AD-A051 147 INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS INC--ETC F/G 9/4
INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY HELD OCTOBER 10-1--ETC(U)
1977 T BERGER, R E BLAHUT

UNCLASSIFIED F49620-77-C-0128
AFOSR-TR-78-0195 NL

1 OF 2
AD A051147

ABSTRACTS OF PAPERS

1977 IEEE

INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY

October 10-14, 1977

CORNELL UNIVERSITY
ITHACA, NY, USA

SPONSORED BY: IEEE INFORMATION THEORY GROUP

CO-SPONSORED BY: UNION RADIO SCIENTIFIQUE INTERNATIONALE

IEEE Catalog Number 77 CH 1277-3 IT

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DDC
This technical report has been reviewed and is approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.
A. D. BLOSE
Technical Information Officer

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR-78-0195	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY held October 10-14, 1977 Cornell University, Ithaca, N.Y. USA	5. TYPE OF REPORT & PERIOD COVERED: Final rept.
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): T. Berger and R. E. Blahut	8. CONTRACT OR GRANT NUMBER(s): F49620-77-C-0128 new
9. PERFORMING ORGANIZATION NAME AND ADDRESS: The Institute of Electrical & Electronics Engineers, Inc. Information Theory Grp, New York, NY 10017	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304 A6
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM Bolling AFB, DC 20332	12. REPORT DATE: 1977
	13. NUMBER OF PAGES: 143
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 150 p.	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This volume contains abstracts of all the papers presented at the 1977 symposium cited.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

181 100

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

1977 IEEE INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY

CORNELL UNIVERSITY
ITHACA, NEW YORK, USA
OCTOBER 10-14, 1977

Sponsored by:
The Institute of Electrical and Electronics Engineers,
Information Theory Group

Co-sponsored by:
Union Radio Scientifique Internationale

Co-Chairmen:
T. Berger and R. E. Blahut

International Advisory Committee:

R. Ahlswede (West Germany)
J. B. Anderson (Canada)
S. Arimoto (Japan)
I. Bar - David (Israel)
P. Bergmans (Belgium)
D. E. Boekee (The Netherlands)
A. B. Carleial (Brazil)
D. L. Cohn (U.S.A.)
I. Csiszar (Hungary)
P. Delsarte (Belgium)
S. Gelfand (U.S.S.R.)
J. Gordon (United Kingdom)
S. J. Halme (Finland)
J. Justesen (Denmark)
G. Longo (Italy)
B. C. Picinbono (France)
J. P. M. Schalkwijk
(The Netherlands)
B. D. Sharma (India)
G. Ungerboeck (Switzerland)
E. C. van der Meulen (Belgium)

Papers:

T. Fine

Program Committee:

V. Chan
K. S. Fu
R. Gray
M. Kaplan
J. Savage
N. Sloane
J. Thomas

Committee Chairmen:

Arrangements: V. Chan
Finance: L. D. Rudolph
Publications: G. M. Groome
Publicity: L. Bahl

IEEE Catalog Number 77CH 1277-3 IT

Library of Congress No. 79-173358

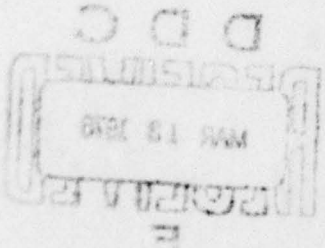
F49620-77-C-0128

ACKNOWLEDGMENTS

We wish to acknowledge the support of

- The National Science Foundation
- The Air Force Office of Scientific Research
- International Business Machines Corporation
- Cornell University

Special acknowledgment goes to IBM for publications and support in the preparation of this volume.



Library of Congress Catalog Card Number: 79-173358

PROGRAM SCHEDULE

SUNDAY, OCTOBER 9, 1977

3:00 PM - 9:00 PM	Registration
7:00 PM	Cocktail Party

MONDAY, OCTOBER 10, 1977

Morning:

8:00 AM	Plenary Session A
9:20 AM	A1: Algebraic Coding Theory A2: Pattern Classification - Statistical Theory A3: Image Processing and Point Processes A4: Multiterminal Channel Coding I A5: Equalization

Afternoon:

2:00 PM	B1: Cryptography B2: Detection I B3: Optical Communications B4: Multiterminal Channel Coding II B5: Noiseless Source Coding

Evening:

8:00 PM	Panel Discussion of Theoretical and Applied Cryptography

TUESDAY, OCTOBER 11, 1977

Morning:

8:00 AM	Plenary Session B
9:20AM	C1: Fast Algorithms (Invited Session) C2: Estimation and Statistical Theory C3: Tree Coding for Channels C4: Rate Distortion Coding C5: Stochastic Processes I

Afternoon:

2:15 PM	Plenary Session C

Evening:

8:00 PM	Cello Concert
9:30 PM	Session on Recent Results

WEDNESDAY, OCTOBER 12, 1977
Morning:

9:00 AM	D1: Communications I
	D2: Pattern Classification - Features
	D3: Topics in Coding
	D4: Universal Source Coding
	D5: Computation I

Afternoon:

2:00 PM	E1: Topics in Algebraic Codes
	E2: Stochastic Processes II
	E3: Speech
	E4: Channel Coding Theory
	E5: Computation II

Evening:

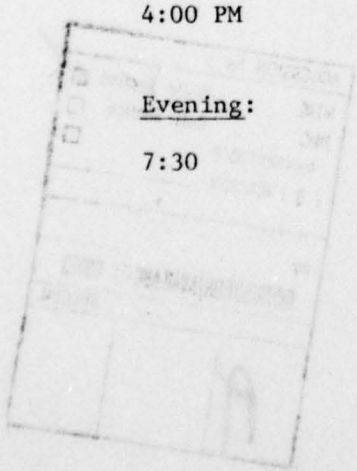
8:00 PM	Discussion: Current Concerns of IEEE

THURSDAY, OCTOBER 13, 1977
Morning:

8:00 AM	Movie: ESP Experiments
9:00 AM	F1: Communications II
	F2: Detection II
	F3: Networks and Queues
	F4: Rate Distortion Theory
	F5: Foundations and Innovations (Special Session)

Afternoon:

1:30 PM	G1: Algebraic Codes
	G2: Identification
	G3: Mathematics in Information Theory (Invited Session)
4:00 PM	Plenary Session D: Shannon Lecture



Evening:

7:30	Banquet

TECHNICAL PROGRAM AND TABLE OF CONTENTS

MONDAY MORNING, OCTOBER 10, 8:00, Uris Hall Auditorium

PLENARY SESSION A

Invited Lecture:

* * * * * *

MONDAY MORNING, OCTOBER 10, 9:20, ILR Conference Center

SESSION A1
ALGEBRAIC CODING THEORY

Chairman, H.F. Mattson Jr., Syracuse University, Syracuse, N.Y. (USA)

SESSION A2
PATTERN CLASSIFICATION - STATISTICAL THEORY

Chairman, K. Fukunaga, Purdue University, West Lafayette, Indiana (USA)

SESSION A3
IMAGE PROCESSING AND POINT PROCESSES

Chairman, Lee D. Davisson, University of Maryland, College Park, Maryland (USA)

*Denotes long papers.

* * * * * *

MONDAY AFTERNOON, OCTOBER 10, 2:00, ILR Conference Center

SESSION B1
CRYPTOGRAPHY

Chairman, Ian F. Blake, University of Waterloo, Waterloo, Ontario (Canada)

*Denotes long papers.

SESSION B2
DETECTION 1

Chairman, Jack Capon, Lincoln Laboratory, Lexington, Massachusetts (USA)

SESSION B3
OPTICAL COMMUNICATIONS

Chairman, Robert S. Kennedy, Massachusetts Institute of Technology, Cambridge, Mass. (USA)

*Denotes long papers.

SESSION B4
MULTITERMINAL CHANNEL CODING II

Chairman, Edward C. van der Meulen, Katholieke Universiteit te Leuven, Heverlee (Belgium)

SESSION B5
NOISELESS SOURCE CODING

Chairman, Demetrios G. Lainiotis, State University of New York at Buffalo, Amherst, NY (USA)

* * * * * *

MONDAY EVENING, OCTOBER 10, 8:00, Phillips Hall 101

Theoretical and Applied Cryptography - Panel Discussion. Organized by N. J. A. Sloane (Bell Labs)

Panel Members: Howard Campaigne (Eastern New Mexico University), Cipher Deavours (Kean College), Horst Feistel (IBM Research), Martin Hellman (Stanford University), Ronald Rivest (MIT), and Aaron Wyner (Bell Labs)

* * * * * *

TUESDAY MORNING, OCTOBER 11, 8:00, Uris Hall Auditorium

PLENARY SESSION B

Invited Lecture:

* * * * * *

TUESDAY MORNING, OCTOBER 11, 9:20, ILR Conference Center

SESSION C1
FAST ALGORITHMS
(INVITED SESSION)

Organizer and Chairman, N. Pippenger, IBM Research, Yorktown Heights, New York (USA)

SESSION C2
ESTIMATION AND STATISTICAL THEORY

Chairman, J. J. Bussgang, Signatron Inc., Lexington, Massachusetts (USA)

SESSION C3
TREE CODING FOR CHANNELS

Chairman, Andrew J. Viterbi, Linkabit Corp., San Diego, California (USA)

SESSION C4
RATE DISTORTION CODING

Chairman, Giuseppe Longo, Universita Degli Studi di Trieste, Trieste (Italy)

SESSION C5
STOCHASTIC PROCESSES I

Chairman, Robert Boorstyn, Polytechnic Institute of N.Y., Brooklyn, NY (USA) and Bell Laboratories, Holmdel, N.J.(USA)

* * * * * *

TUESDAY AFTERNOON, OCTOBER 11, 2:15, Phillips Hall 101

PLENARY SESSION C

Invited Lectures:

* * * * * *

TUESDAY EVENING, OCTOBER 11

8:00 Cello Concert by Nancy Green
Barnes Hall Auditorium

* * * * * *

9:30 Session on Recent Results and Open Problems
Phillips Hall 101

Chairman, David L. Cohn, University of Notre Dame, Notre Dame, Indiana (USA)

Titles and speakers to be arranged at the Symposium.

* * * * * *

WEDNESDAY MORNING, OCTOBER 12, 9:00, ILR Conference Center

SESSION D1
COMMUNICATIONS I

Chairman, Jack Keil Wolf, University of Massachusetts, Amherst, Massachusetts (USA)

SESSION D3
TOPICS IN CODING

SESSION D4
UNIVERSAL SOURCE CODING

*Denotes long papers.

SESSION D5
COMPUTATION I

Chairman, Judea Pearl, University of California, Los Angeles, California (USA)

* * * * * *

WEDNESDAY AFTERNOON, OCTOBER 12, 2:00, ILR Conference Center

SESSION E1
TOPICS IN ALGEBRAIC CODES

Chairman, R. T. Chien, University of Illinois at Urbana-Champaign, Urbana, Illinois (USA)

*Denotes long papers.

* * * * * *

WEDNESDAY EVENING, OCTOBER 12, 8:00, Phillips Hall 101

"Current Concerns of IEEE: A Discussion" by Irwin Feerst and Ivan Getting, Candidates for President of IEEE. Moderated by F. Jelinek . . . . . .

* * * * * *

THURSDAY MORNING, OCTOBER 13, 8:00, ILR Conference Center

Movie on ESP Experments

* * * * * *

THURSDAY MORNING, OCTOBER 13, 9:00, ILR Conference Center

SESSION F1
COMMUNICATIONS II

Chairman, Robert Price, Sperry Research Center, Sudbury, Massachusetts (USA)

SESSION F2
DETECTION II

Chairman, Michael B. Pursley, University of Illinois at Urbana - Champaign, Urbana, Illinois (USA)

SESSION F3
NETWORK AND QUEUES

Chairman, David Slepian, Bell Telephone Laboratories, Murray Hill, N.J. (USA)

SESSION F5
FOUNDATIONS AND INNOVATIONS
(Special Session)

Chairman, Terrence Fine, Cornell University, Ithaca, N.Y. (USA)

* * * * * *

THURSDAY AFTERNOON, OCTOBER 13, 1:30, ILR Conference Center

SESSION G1
ALGEBRAIC CODES

Chairman, I. S. Reed, University of Southern California, Los Angeles, California (USA)

†Denotes invited paper.

SESSION G2
IDENTIFICATION

Chairman, Emanuel Parzen, State University of New York at Buffalo, Amherst, New York (USA)

SESSION G3
MATHEMATICS IN INFORMATION THEORY
(Invited Session)

Organizer and Chairman, Paul Shields, University of Toledo, Toledo, Ohio (USA)

* * * * * *

*Denotes long papers.

THURSDAY AFTERNOON, OCTOBER 13, 4:00, Phillips Hall 101

PLENARY SESSION D

SHANNON LECTURE - Peter Elias, Massachusetts Institute of Technology, Cambridge, Mass., USA

* * * * * *

THURSDAY EVENING, OCTOBER 13, 7:30

Banquet

* * * * * *

ABSTRACTS

PLENARY SESSION A

MULTIPLE TERMINAL CHANNEL THEORY - Thomas M. Cover (Information Systems Laboratory, Stanford University, Stanford, CA 94305). A unified presentation of new or recent results in multiple user channel theory will be given. In particular the duality between the multiple access channel capacity (Ahlswede, Liao) and multiple source compression capacity (Slepian, Wolf) will be developed. Repeated use of the asymptotic equipartition property will be seen to yield simple proofs for achievable rate regions for certain broadcast channels, multiple access channels, two-way channels, relay channels, and the above-mentioned channels with feedback.

At this time no theoretical evidence contradicts the hope that we shall eventually have an information - theoretic characterization of the capacity region for an arbitrary information network with m senders, n receivers, and arbitrary feedback loops.

SESSION A1

SOME RESULTS ON A GENERAL CLASS OF QUASICYCLIC CODES, Pierre A. von Kaenel (Department of Mathematics and Computer Science, The University of Nebraska at Omaha, Omaha, Nebraska 68101). A general class of quasicyclic (Q-C) codes which contains many of the special cases investigated in the literature such as (2n,n) Q-C codes will be defined. An algebraic method for designing the subclass of these codes having length 2 (2 blocks of n-tuples) will then be presented. Using this method, the first rows of the circulant matrices contained in the generator matrix of a Q-C code in this subclass can be selected in such a way that the lower bound on minimum distance is considerably improved over that of an arbitrary Q-C code of the same size. Chen et al. have found general lower bounds on minimum distance for (2n,n) Q-C codes; however our designing method can improve this bound by as much as a factor of two. Other classes of Q-C codes can be designed, and some codes can be found which compare favorably with the best known linear codes of that size. For example, a (42, 11) Q-C code can be constructed with a minimum distance d bounded by $14 \leqslant d \leqslant 16$. The best known (42,11) linear code has $d = 15$.

SOME RESULTS ON THE STRUCTURE OF GOPPA CODES, C.L. Chen (D18/707, P.O. Box 390, IBM Corporation, Poughkeepsie, New York 12602). A relationship between Goppa codes and cyclic codes is established. Goppa codes are identified as shortened cyclic codes. Irreducible Goppa codes are classified into equivalence classes to reduce the complexity of the weight enumeration of the codes. The location points of a code word of an irreducible Goppa code are related to the points of a finite projective geometry. The irreducible Goppa codes extended with an overall parity check are investigated. All binary extended irreducible codes of a fixed length with minimum distance 8 are shown to have the same weight distribution.

AN ENUMERATION OF SELF-DUAL CODES, J.H. Conway (University of Cambridge, Cambridge, England) and Vera Pless (University of Illinois at Chicago Circle, Chicago, Illinois 60680). A linear code C is called self-dual if $C = C^{\perp}$. Self-dual codes are a practical and interesting class of codes because they include some of the best algebraic codes like the Golay codes, the quadratic residue codes, and the symmetry codes and also because more is known theoretically about them, i.e., their weight distributions are combinations of Gleason polynomials. Further, it is known that they satisfy a Varshamov-Gilbert bound.

A binary self-dual code must have all weights even. If in addition all weights are divisible by 4, the code is called doubly-even. The Golay (24,12) code is doubly-even as is the Hamming (8,4) code. This paper completely classifies all (32,16) double-even codes.

It is shown that of the $2 \cdot 3 \cdot 5 \cdot \cdots (2^{14}+1)$ self-dual doubly-even codes there are 85 inequivalent such codes, five of which have minimum weight 8, three new codes and the two known codes: a quadratic residue code and a Reed-Muller code. For each code in an equivalence class, a canonical basis is given, as is the entire group of the code and its weight distribution. From this list we are able to identify all self-dual (n,n/2) codes for $n < 32$. The data for $n = 30$, 28, and 26 is new. Further, we show that for the next case (40,20), the number of inequivalent double-even codes is greater than 17,000 so that it is not possible to classify them with present or forseeable methods.

CODES OVER GF(4) AND COMPLEX LATTICES, N.J.A. Sloane (Bell Labs, Murray Hill, N.J.). This paper studies the relationship between error-correcting codes over GF(4) and complex lattices. The theta-functions of self-dual lattices are characterized. Two general methods are presented for constructing lattices from codes. Several examples are given, including a new lattice sphere-packing in R^{36}.

CODES WITH REGULAR GAPS IN THEIR WEIGHT DISTRIBUTION, V.K. Bhargava, (Department of Electrical Engineering, Concordia University, Montreal, Canada H3G 1M8). Consider (mn,m) rate 1/n quasi-cyclic codes whose generator matrices in the partitioned form are given by $G = [I_n, C_1, C_2, \ldots, C_{n-1}]$. By choosing the circulant matrices C_i's judiciously, we have discovered a number of new rate 1/n quasi-cyclic codes with regular gaps in their weight distribution. These, among others, include a (26,13) $d = 5$ code with weights $\equiv 1$ or 0 (mod 4), a (52,13) $d = 15$ code with weights $\equiv 3$ or 0 (mod 4), a (54,18) $d = 12$ code with weight $\equiv 0$ (mod 4) and a family of (3p,p) codes with weights $\equiv 3$ or 0 (mod 4). We conclude with a number of rate 1/2 and rate 1/3 quasi-cyclic codes showing regular gaps in their weight distribution.

A NEW CLASS OF ASYMPTOTICALLY GOOD CODES BEYOND THE ZYABLOV BOUND, Yasuo Sugiyama (Mitsubishi Electric Corporation, 80 Nakano, Minamishimizu, Amagasaki, Hyogo 661, Japan), Masao Kasahara (Department of Communication Engineering, Osaka University, Yamada-Kami, Suita, Osaka 565, Japan), Shigeichi Hirasawa (Mitsubishi Electric Corporation, 1-1-2 Wadasakicho, Hyogoku, Kobe, Hyogo 652, Japan) and Toshihiko Namekawa (Department of Communication Engineering, Osaka University, Yamada-Kami, Suita, Osaka 565, Japan). We construct a new class of asymptotically good codes. The asymptotically good codes are constructed by iteratively superimposing Justesen codes on a Justesen code. The codes can be regarded as members of a class of generalized concatenated codes. For any rate r, $0 < r < 1$, the ratio d/n of minimum distance d to length n of the codes satisfies the relation.

$$\liminf_{n\to\infty} (d/n) \geq \frac{(1-2^{-\ell})-r}{\sum_{j=1}^{\ell} [1/(\theta^{(j)}\, 2^{j})]}\; H^{-1}(1/2)$$

where ℓ is a positive integer, $H(\cdot)$ is the binary entropy function, and $\theta^{(j)} = H^{-1}(1/2^{j})/H^{-1}(1/2)$ for $j=1,2,\ldots,\ell$. The lower bound on the ratio d/n for the new codes lies above the Zyablov bound for rate r, $0.31 < r < 1$, while the lower bound on the ratio d/n for all asymptotically good codes so far known to us lies below or on the Zyablov bound.

COMPUTER DETERMINATION OF THE TRUE MINIMUM DISTANCE OF ALL BINARY CYCLIC CODES OF ODD LENGTHS FROM 69 to 99, G. Promhouse and S.E. Tavares (Department of Electrical Engineering, Queen's University, Kingston, Ontario, Canada). A computer search has been made to determine the true minimum distance d for all binary cyclic codes having odd lengths n in the range $69 \leq n \leq 99$. Using an algorithm originally developed by C.L. Chen, the generator matrix of each (n,k) binary cyclic code was determined and put in systematic form G. All possible code words obtained by summing the rows of G, i at a time, $i = 1,2,\ldots,v$, were examined and the minimum distance d_v of this set was recorded. Then $d = d_v$ whenever $v > \{ (d_v-1)k/n \} - 1$. Like Chen, known equivalences among cyclic codes were taken into account and only one code from each equivalence class was listed. Let g(x) divide (x^n-1), where (x-1) is not a factor of g(x). Then the minimum distance of the codes generated by g(x), (x-1) g(x) and their respective duals were listed together. For each such pair of codes in the table, the value of v for which a code word of minimum weight first appeared is listed. The codes found were compared with the list of best codes tabulated by Sloane. Many good cyclic codes have been found. Among the best (n,k,d) cyclic codes found are the following: (73,36,16), (73,27,20), (85,20,28), (89.56,11), (91,51, 14), (91,28,24), (93,23,29).

SESSION A1

LINEAR RECURSIVE CODES, Michael Willett (Mathematics Department, University of North Carolina at Greensboro). Let C be an (N,K) cyclic code over $F = GF(p^e)$ with parity-check polynomial $h(x) = h_1(x)h_2(x)\ldots h_t(x)$, where the $h_i(x)$ are distinct and irreducible over F. It is well-known that C equals G(h), the vector space of all N-tuples $u = (u_n)_o^{N-1}$ satisfying $h(E)u_n = 0$ with $Eu_n = u_{n+1}$ and all subscripts reduced modulo N. Linear recursion theory is used to show that it is sometimes possible to choose $s = (k_1, k_2, \ldots, k_t)$, $k_i \geq 1$, $K_s = \Sigma k_i \deg(h_i) < N$, in such a way that the (N,K_s) (non-cyclic) linear code $G(h_s)$, $h_s(x) = h_1^{k_1}(x)\, h_2^{k_2}(x)\ldots h_t^{k_t}(x)$, has the same minimum distance as the subcode G(h).

CONVERGENCE BOUNDS OF BAYES LEARNING ESTIMATORS, Ary Kiselstein and Amnon Hammer (Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa, Israel). The paper analyzes the convergence of signal estimation by means of Bayes learning. The true signal m_t is assumed as deterministic, fixed, and embedded in additive Gaussian noise $N(0,\sigma^2)$. The quality of convergence depends on the mean m_o, and variance of ϕ_o^2 of the initially assumed signal probability density. The estimator performance is evaluated by means of the estimate variance $\overline{\varepsilon_n^2}$. The value of $\overline{\varepsilon_n^2}$ is of course minimal if $m_o = m_t$. When m_o is chosen as $m_o = m_t \pm \beta$, the optimal choice for ϕ_o^2 becomes $\phi_o^2 = \beta^2$, yielding a bound for $\overline{\varepsilon_n^2}$ as

$$\overline{\varepsilon_n^2}\Big|_{\beta=\text{const.}} \geq \frac{\sigma^2\beta^2}{\sigma^2 + n\beta^2}$$

The value of $\overline{\varepsilon_n^2}$ does not converge down monotonically, with increase of n, only if $\phi_o^2 > 2\beta^2$.

The number of learning intervals required to decrease $\overline{\varepsilon_n^2}$ to a fraction δ of its initial value is given by $\frac{1}{\delta} - \frac{\sigma^2}{\beta^2}(1 - \frac{1}{\delta})$.

It is shown that the Cramer-Rao bound exceeds the Bayes learning estimator bound when $\phi_o^2 = \beta^2$. The paper is concluded with numerical examples.

OPTIMAL BAYES DECISION RULES IN MULTI CATEGORY CLASSIFICATION PROBLEMS WITH TIME SERIES, R.L. Kashyap (School of Electrical Engineering, Purdue University, West Lafayette, Indiana 47907). We consider the problem of classifying a given time series $Z_N = \{y(1), \ldots, y(N)\}$ into one of r classes C_i, $i=1,\ldots,r$. The stochastic process $y(\cdot)$ is assumed to obey an autoregressive structure involving a parameter vector $\underset{\sim}{\theta}$, whose probability density $p(\underset{\sim}{\theta}|C_i)$ depends on the class to which Z or $y(\cdot)$ belongs. Assuming appropriate expressions for $p(\underset{\sim}{\theta}|C_i)$, we show that the probability density of Z_N characterizing each class, namely $p(Z_N|C_i)$, possesses a vector $\overline{\underset{\sim}{\theta}}$ of sufficient statistics i.e., all the information about Z_N needed for the discrimination between the various classes is contained in the vector $\overline{\underset{\sim}{\theta}} = (\overline{\theta}_1(Z_N), \ldots, \theta_{m+1}(Z_N))^T$, where the functions $\overline{\theta}_i(Z_N)$, $i=1,\ldots,m+1$ have the same structure for all N. Thus the best possible feature set for the problem is $\overline{\underset{\sim}{\theta}}$. We deduce the optimal decision rule which minimizes the average probability of error. We will compare the optimal features and the corresponding optimal decision rule with other feature sets and decision rules mentioned in the literature.

ON AN ADMISSIBLE LINEAR DISCRIMINATION RULE FOR MULTICLASS PATTERN RECOGNITION, Dimitri Kazakos and Basile Dimitriadis (Department of Electrical Engineering, State University of New York at Buffalo, Amherst, New York 14260). Let X be an n-dimensional observation which we wish to classify according to the linear rule:

$$\text{decide } H_i \text{ if } C_{i-1} < X^t V \leq C_i,\ i = 1,\ldots, m.$$

Let (M_i, R_i) be the mean and covariance matrix of X under hypothesis H_i, $i = 1\ldots, m$, $m \geq 3$.

Let $P_{di}(D)$ be the probability of correct classification under H_i, where:

$$D = \{V, -\infty = C_o < C_1 <\ldots< C_{m-1} < C_m = +\infty\}$$

A rule D* is said to be admissible if there is no other rule D_1 for which

$$P_{di}(D_1) \geq P_{di}(D^*),\ i = 1,\ldots, m$$

In the present paper we show that, under mild conditions on the probability density functions $\{f(X|H_i)\}$, the class of admissible rules D* belongs to the set:

$$\{V, [aR_i + bR_j + cR_k] V = d(M_i - M_j) + e(M_j - M_k),\ i,j,k\varepsilon\{1,\ldots, m\},\ a,b,c \geq 0\}$$

(the constants $\{C_i\}$ are given in terms of V for the admissible class).

OPTIMIZATION OF ADD-ON SIGNALS BY MEANS OF A MODIFIED TRAINING ALGORITHM FOR LINEAR CLASSIFIERS, Wolfram Szepanski (Institute fur Elektrische Nachrichtentechnik, Rheinisch-Westfälische Technische Hochschule Aachen, D-5100 Aachen, Alte Maastrichter Str. 23, Germany). In add-on data transmission for TV channels, a data signal is added to the active parts of the video signal in order to transmit additional information.

Compatibility requirements and their impact on the selection of suitable data signals are discussed. To make maximum use of the correlation within the video signal, the data and the video signals are considered to be functions of three variables x, y, and t.

Since correlation reception of these three-dimensional signals corresponds to a linear classification problem in pattern recognition, a modified training procedure for linear classifiers is used to optimize the data signals for minimum error probability.

The results of the optimization procedure are presented together with achievable data rates and error probabilities for this type of data transmission over TV channels.

QUANTIZATION COMPLEXITY AND TRAINING SAMPLE SIZE IN PATTERN RECOGNITION, Dimitri Kazakos (Electrical Engineering Department, State University of New York at Buffalo, Amherst, New York 14260). For the k-hypothesis detection problem, it is shown that among the k-classes of probability density functions with m fixed quantiles the histograms achieve the least favorable performance, as it is measured by the probability of correct detection and Chernoff distance.

We assume, then, that the m cell probabilities are estimated using n training samples per class. Using the estimated cell probabilities, new observations are processed. A distribution-free upper bound to the probability of ε-deviation between the actual probability of correct detection and the theoretical (known quantities) probability is derived, given as a function of (m,n,ε,k,u_o), where u_o is uniform upper bound to the true class densities. The bound converges exponentially to zero as $n \to \infty$. Exponential convergence is maintained by choosing $m = n^{\alpha}$, $0 < \alpha < 1$. Hence, the rule $m = n^{\alpha}$ answers the long standing question of relating m and n in a distribution-free manner. The question of optimal choice of α is also discussed.

ERROR ESTIMATION AND SAMPLE DESIGN, Stephen J. Whitsitt (Data Systems Department, TRW DSSG, One Space Park, Bldg. 90, Room 2200, Redondo Beach, California 90278) and David A. Landgrebe (Laboratory for Applications of Remote Sensing, Purdue Industrial Research Park, 1220 Potter Drive, W. Lafayette, Indiana 47906). An estimator for error which employs both posterior probabilities and sample stratification by class is developed as an extension, and completion of a family of count and posterior estimators. The effectiveness of this estimator is demonstrated using experiments involving normal patterns. Applications in testing fixed classifiers are discussed.

DISTRIBUTION-FREE PROBABILITY INEQUALITIES FOR THE DELETED AND THE HOLDOUT ERROR ESTIMATES, L.P. Devroye and T.J. Wagner (University of Texas, Austin, Texas 78712). Let $(X_1,\theta_1),\ldots,(X_n,\theta_n)$ be a sequence of independent $\mathbb{R}^d \times \mathbb{R}$-valued random vectors distributed as (X,θ). A discrimination rule $\hat{\theta}$, a real-valued function of x and the data, induces a probability of error $L_n = P\{\hat{\theta} \neq \theta \mid (X_1,\theta_1),\ldots,(X_n,\theta_n)\}$. To estimate L_n from $\hat{\theta}$ and the data, the resubstitution, deleted and holdout estimates can be used.

For rules that use X and some or all of the k nearest neighbors to X (in an arbitrary way), upper bounds are derived for $P\{|\hat{L}_n - \hat{L}_n| > \varepsilon\}$ where $\hat{L}_n$ is one of the said estimates. The bounds are distribution-free, and depend only upon n, k and ε. Exponential bounds exist that also depend upon d. The inequalities obtained for the deleted estimate are the tightest.

USE OF MAJORITY-VOTE IN n-SAMPLE DECISION MAKING, Sargur N. Srihari (Computer Science Department, Wayne State University, Detroit, MI 48202). The use of context in pattern recognition is achieved by the application of compound decision rules. The n-sample decision problem is a special case of the compound decision problem where each of n samples is known to belong to the same but unknown state of nature. A natural method of utilizing a simple decision rule in determining the true state of nature is to classify each sample individually and choose the class represented most often among n decisions. This is the majority-vote strategy which possesses the following important properties.

1. It can be formulated as a sequential procedure.

2. Its n-sample error rate can be evaluated as a function of the class-conditional error probabilities of the simple decision rule.

3. Its error rate for two equally likely classes, operating with a decision rule of specified average misclassification probability, increases monotonically with classification bias, β, of the decision rule; where β is a decision rule parameter defined as the difference between class-conditional error probabilities for the problem.

The paper describes the application of these results in a practical nonparametric pattern classification problem where β is shown to be a useful criterion for decision rule or feature set evaluation.

SESSION A3

SAMPLE FUNCTION REGULARITY OF SHOT PROCESSES, Robert Lugannani (Department of Applied Physics and Information Science, University of California, San Diego, La Jolla, California 92093). This paper is concerned with continuity and differentiability of the sample functions of shot processes. A knowledge of these properties is required in applications in which it is desired to apply classical analysis techniques to the individual sample functions of the process. Sufficient conditions are given for the sample functions to be continuous and to be continuously differentiable and it is shown that these conditions are often satisfied in practice. The proof is based on a new bound for the probability of large deviations of the process. This bound has considerable independent interest and its applicability is not limited to the study of sample function properties. With the exception of a few special cases, convenient expressions for the amplitude distribution of a shot process are not known, and in these instances the bound provides valuable information concerning its behavior.

SPACE-TIME CODING FOR IMPROVED OPTICAL LINE-SCAN IMAGING, Stanley R. Robinson (School of Engineering, Air Force Institute of Technology, Department of Electrical Engineering, Wright-Patterson AFB, Ohio 45433). An interesting form of active, high resolution imaging is that obtained by an optical line scan system. In such a system the use of a modulated laser would make possible slant-range (or height) as well as reflectivity measurements for each resolution cell. This paper considers the problem of efficiently using multiple solid state laser sources to improve the range/reflectivity estimation performance of such a system. Only direct detection receivers and power (intensity) modulated sources are considered viable candidates.

We demonstrate that multiple sources can be efficiently used if each laser is focused to a spatial resolution cell disjoint from all other sources and if each is modulated with a waveform (or code word) which is essentially temporally orthogonal from all others. Thus each spatial resolution cell is coded with a unique temporally modulated signal which can be decoded by the receiver.

The detector used in the receiver is assumed to be an ideal photon detector, i.e., its output is modeled as a conditional Poisson point process. The maximum likihood (ML) range/reflectance estimator structures are presented; an indication of the variance of the localized estimate is obtained through the Cramer-Rao lower bound. It is shown that both variances will improve by as much as $\frac{1}{n}$ where n is the number of sources. Motivated by the variance of the range estimate, we select an ON-OFF modulation scheme whereby the laser is modulated by a Pseudonoise (PN) sequence whose period is chosen to be the dwell time. We conclude that acceptable estimation performance in both range and reflectance in each cell can be attained with such a multiple source-modulation format.

SESSION A3

INTERFRAME CODING OF VIDEOTELEPHONE PICTURES USING MOVEMENT COMPENSATION - RECENT SIMULATIONS, B.G. Haskell and D.K. Sharma (Bell Laboratories, Holmdel, New Jersy 07733). Recent results are reported for adaptive interframe coding which takes advantage of the fact that moving areas in one television frame are often very similar to corresponding areas in the previous frame except for a linear translation. Minimum M.S.E. linear prediction and interpolative prediction using the Limb Velocity-Measure were employed to code 30 successive videotelephone frames containing movement. Entropies ranged between 1 and 2 bits per moving-area pel. A videotape of coded pictures will be shown.

NEAR-OPTIMAL ESTIMATION-DETECTION SCHEME FOR POISSON-DRIVEN PROCESSES, S. Au and A.H. Haddad (Department of Electrical Engineering and Coordinated Science Laboratory, University of Illinois, Urbana, Illinois 61801). A near-optimal scheme for the estimation of signals with linear models and a Poisson process input is proposed. The scheme is nonlinear and is based on the detection of the input jumps in small subintervals which are then used to estimate the signal. The scheme is suitable for the case of small incident rate λ of the Poisson input. The performance is compared to that of optimal linear filters which are optimal for such processes only for large λ. The results indicate that there exists a rate λ^* such that the proposed scheme performs better in the MMSE sense than the linear scheme for $\lambda < \lambda^*$.

PERFORMANCE FOR BINARY DECISION PROBLEMS WITH POINT PROCESS OBSERVATIONS, J.L. Hibey (Dynamics Research Corporation, Systems Division, 60 Concord Street, Wilmington, MA 01887), D.L. Snyder and J.H. van Schuppen (Washington University, St. Louis, MO 63130). We are concerned with the evaluation of error-probability bounds for binary detection problems involving point process observations. These bounds are of interest because the computation of the exact probabilities of error is usually mathematically intractable. The basic approach we employ is the application of martingale theory to detection and estimation problems.

The decision strategy we employ is the generalized likelihood-ratio test, and the bounding technique is due to Chernoff. A measure transformation technique allows us to obtain an expression for the Chernoff bound in terms of an expectation of a multiplicative functional of the conditional mean signal (rate process) estimates. If the processes involved are Markovian, we are then able to represent the above expression as a solution to a backward Kolmogorov equation.

The above procedure is essentially repeated when the optimal estimates are replaced by suboptimal estimates. These results are particularly important because the optimal filtering equations cannot be solved exactly, numerical solutions are computationally infeasible, the optimal estimates are non-Markovian, and explicit systems for generating the exact estimates are known only in a few cases.

We then illustrate our procedure with examples motivated by problems of of interest in optical communications.

EXPERIMENTAL PERFORMANCE OF POINT PROCESS ESTIMATORS OF OPTICAL PULSE DELAY, F. Davidson and L. Stephens (Department of Electrical Engineering, The Johns Hopkins University, Baltimore, Maryland 21218). A series of experiments was performed to evaluate the performance of minimum mean square error (MMSE) estimates, $\hat{\tau}$, of the delay, τ, of Gaussian and rectangular shaped optical pulses from observed photoelectron emission times. In the absence of backgound radiation, the mean square error for rectangular pulses (13 μsec wide, rise time of 0.5 μsec) was found to decrease as Q^{-2} as predicted by Bar-David, where Q is the optical pulse energy. Gaussian shaped pulses of equal peak intensity and identical Q were found to have the predicted Q^{-1} dependence in the mean square error of estimates of τ.

MMSE estimates in the presence of background radiation of constant intensity λ_0 were found from numerical solutions to equations of evolution for the posterior conditional density of τ found by Snyder. At equal pulse energies and peak signal-to-noise ratios, λ_s/λ_0, the rectangular pulses were found to give superior performance over the Gaussian shaped pulses for values of λ_s/λ_0 in excess of 10. As the ratio λ_s/λ_0 approached unity, the mean square errors in $\hat{\tau}$ for the two differently shaped pulses were observed to become nearly equal. At high signal to noise ratios (i.e., > 10), the mean square error in $\hat{\tau}$ for the Gaussian shaped pulse was found to be nearly equal to the lower bound found by Snyder.

IMAGE MODELING WITH APPLICATION TO MEASUREMENT, J.W. Burnett (TRW Systems, 1 Space Park, Redondo Beach, CA 90278), and T.S. Huang (School of Electrical Engineering, Purdue University, West Lafayette, Indiana 47907). We develop a fast algorithm for pulse width estimation from blurred and nonlinear observation in the presence of signal-dependent noise. The problem is motivated by the need for accurate measurements from remotely sensed photographs. Ideally, reflected light intensity from an object and its background will be approximately piecewise constant with discontinuities at the edges.

However, after the reflected light has passed through the imaging system, the edges have been blurred. Film then responds nonlinearly to the incident light and adds signal-dependent noise.

The problem is approached by modeling the edge signal (reflected light intensity) as a discrete-position, finite-state Markov process. Sample functions of such a process are graphically represented by a path through a trellis. Blurred versions of these signals are similarly represented. By assigning a cost to each branch of the trellis, a maximum a posteriori probability sequence estimate of the signal is computed by finding the minimum cost path through the trellis. The Viterbi algorithm is introduced as an efficient means of finding the minimum cost path through the trellis.

The algorithm is applied to the measurement of a road in an aerial photo taken at an altitude of 5000 feet. The resulting width estimate is accurate to within a few inches.

ENCODING OF IMAGES ON A TWO-COMPONENT SOURCE MODEL, Johnson K. Yan and David J. Sakrison (Department of Electrical Engineering and Computer Science and the Electronics Research Laboratory, University of California, Berkeley, California 94720). Setting aside the issue of complexity, transform codes seem to perform the most efficiently for transmission of grey-scale still images. For large raster sizes, a Fourier transform code is optimum under the assumption of a Gaussian Source and weighted-square-error distortion criterion. A weighted-square-error distortion criterion on log-intensity does not represent too badly the criterion of a viewer's visual system when errors are just barely visible. However, sample images look nothing like sample fields from a Gaussian random field. A better code can thus result from a better modeling of the source. We model the source by breaking a sample field up into two components: a "discontinuous component" representing abrupt changes due to distinct objects, and a continuous component representing shading and texture. The discontinuous component is represented as a ruled surface, horizontal scan lines across it consisting of straight line segments. The breakpoints of these segments are quantized and differentially encoded. The continuous component is the remainder after subtracting the discontinuous component from the original; it has the behavior of a Gaussian field and is encoded by a transform code. Simulations reveal considerable improvements over conventional transform coding of the whole image.

RATE-DISTORTION WITH A FULLY INFORMED DECODER AND PARTIALLY INFORMED ENCODER, Toby Berger and Ming U. Chang (School of Electrical Engineering, Cornell University, Ithaca, N.Y. 14853). Consider the following generalization of the rate-distorion problem of Wyner and Ziv. The discrete memoryless source (dms) $\{X_k\}$ is to be encoded for transmission to a destination at a rate of R_x bits per second (bps). Side information is available at said destination in the form of another dms, $\{Y_k\}$. The joint distribution of X_k and Y_k, which does not depend on k, is denoted by P. The generalizaton is that information about $\{Y_k\}$ may be supplied to the encoder at a rate of R bps; Wyner and Ziv treated the special case R=0.

Let R_x (D,R) denote the least value of R_x that sufficies to permit reconstruction of $\{X_k\}$ at the destination with fidelity D as assessed by means of a memoryless distortion measure, $\rho(x,\hat{x})$. By proving a coding theorem we show that $R_x(D,R) \leq \inf I(X;Z|Y,W)$, where (X,Y) is distributed according to P and the infimum is over all discrete random variables W and Z that satisfy the following four conditions:

(i) X,Y,W is a Markov chain

(ii) Y, (X,W), Z is a Markov chain

(iii) $I(Y;W|X) \leq R$

(iv) there exists a function f such that $\hat{X} \triangleq f(Z,Y)$ satisfies $E\rho(X,\hat{X}) \leq D$.

We also show that $R_x(D,R) \geq \inf I(X;Z|Y,W)$ if the infimum is taken over all discrete W and Z that need satisfy only (i), (iii), and (iv) above. Both our upper and our lower bound are non-trivial in the sense that, for all $R\varepsilon(0,H(Y|X))$, the former is less than or equal to the Wyner-Ziv rate-distortion function and the latter is greater than or equal to the conditional rate-distortion function. For R = 0 the upper bound reduces to the answer provided by Wyner and Ziv, and for $R \geq H(Y|X)$ the upper and the lower bound both equal the conditional rate-distortion function $R_{X|Y}(D)$.

MULTITERMINAL SOURCE CODING, Toby Berger and Sui-Yin Tung (School of Electrical Engineering, Cornell University, Ithaca, N.Y. 14853). Let $\{\underline{X}_k = (X_{1k},\ldots X_{nk}),\ k = 0, \pm 1, \pm 2,\ldots\}$ denote a sequence of independent drawings from a known n-variate distribution $P_{\underline{X}}$. The sources $\{X_{ik},\ k = 0, \pm 1, \pm 2,\ldots\}$, $1 \leq i \leq n$, can be encoded at respective rates R_i for transmission to a common destination. The rate vector $\underline{R} = (R_1,\ldots,R_n)$ is $\underline{D} = (D_1,\ldots,D_n)$ - admissible if, on the basis of all n encoder outputs delivered to the destination, it is possible to reconstruct $\{X_{ik}\}$ there with average distortion D_i or less as assessed by means of the memoryless distortion measure $\rho_i(x,\hat{x})$, $1 \leq i \leq n$. Let $R(\underline{D})$ denote the region of all $\underline{D}$-admissible rate vectors.

We have established conditions sufficient to ensure that $\underline{R} \in R(\underline{D})$; we describe them here only for the case of n=2. Let $\underline{X} = (X_1, X_2)$ be distributed according to $P_{\underline{X}}$. Let $R_{in}(\underline{D})$ denote the set of all $\underline{R} = (R_1, R_2)$ such that there exists a random vector $\underline{Y} = (Y_1, Y_2)$ that satisfies

(i) $I(Y_1; (X_2, Y_2) | X_1) = I(Y_2; (X_1, Y_1) | X_2) = 0$

(ii) $I(X_1; Y_1 | Y_2) \leq R_1$, $I(X_2; Y_2 | Y_1) \leq R_2$, $I(X_1; Y_1) + I(X_2; Y_2 | Y_1) \leq R_1 + R_2$

(iii) there exists functions f_1 and f_2 such that, where $\hat{X}_i = f_i(\underline{Y})$, we have $E\, \rho_i(X_i, \hat{X}_i) \leq D_i$, i=1,2.

Then $R(\underline{D}) \supseteq R_{in}(\underline{D})$.

We also show that $R(\underline{D}) \subseteq R_{out}(\underline{D})$, where $R_{out}(\underline{D})$ is defined in the same manner as is $R_{in}(\underline{D})$ except that (i) and (ii) are replaced by

(i)' $I(Y_1; X_2 | X_1) = I(Y_2; X_1 | X_2) = 0$

(ii)' $I(\underline{X}; Y_1 | Y_2) \leq R_1$, $I(\underline{X}; Y_2 | Y_1) \leq R_2$, $I(\underline{X}; \underline{Y}) \leq R_1 + R_2$

In general, $R_{in}(\underline{D})$ is a strict subset of $R_{out}(\underline{D})$.

For a bivariate Gaussian $P_{\underline{X}}$ with parameters σ_1^2, σ_2^2, and ρ, let

$$D(\underline{D}) = \{(d_1, d_2): d_i \leq D_i/\sigma_1^2,\ d_j = D_j/\sigma_j^2,\ i,j \in \{1,2\},\ i \neq j\}.$$

Then $\underline{R} \in R_{in}(\underline{D})$ iff there exists $(d_1, d_2) \in D(\underline{D})$ such that, where

$$\beta = 1 + \sqrt{1 + 4\rho^2 d_1 d_2 (1-\rho^2)^{-2}},$$

$$R_1 + R_2 \geq \frac{1}{2} \log \left[\frac{(1-\rho^2)\beta}{2 d_1 d_2} \right]$$

$$R_i \geq \frac{1}{2} \log \left[\frac{(1-\rho^2)^2 \beta}{(1-\rho^2)\beta d_i - 2\rho^2 d_1 d_2} \right], \quad i \in \{1,2\}$$

This result is extended to jointly stationary Gaussian random sequences.

CODING RELATIVE TO FIDELITY CRITERIA FOR CORRELATED SOURCES WITH MEMORY, Jim K. Omura (System Science Department, University of California, Los Angeles, CA) and Kim B. Housewright (Systems Science Department, University of California, Los Angeles, CA and Hughes Aircraft Company, Fullerton, CA). Recent research efforts have provided partial characterizations of (i.e., "inner" and "outer" bounds on) the rate-distortion surfaces of multiterminal systems involving memoryless sources. In this short paper we describe an extension of those findings to systems involving sources with ergodic memory structure. We obtain an "inner" bound which is cosmetically similar to the result for the memoryless case, but is defined in terms of regions associated with a constrained set of stochastic processes. In fact, our bound compares to the previous results in exactly the same manner that the "process definition" of the rate distortion function compares to the memoryless result for a single source-single user system. We have also developed similar bounds for more general networks, but the complexity of the requisite notation makes them unsuitable for discussion.

CODE CONSTRUCTION FOR THE T-USER BINARY ADDER CHANNEL, S.C. Chang (Electrical Engineering Department, University of Hawaii, Honolulu, HI 96822) and E.J. Weldon, Jr. (Electrical Engineering Department, University of Hawaii, Honolulu, HI 96822, and Adtech, Inc., P.O. Box 10415, Honolulu, HI 96816). Coding schemes for the discrete memoryless T-user adder channel are investigated in this paper.

First, the capacity region of the noiseless T-user adder channel, and the maximal achievable rate for T-user uniquely decodable codes are derived. Second, code constructions are presented, including a class of T-user uniquely decodable codes with rates asymptotically equal to the maximal achievable value. A decoding algorithm for these codes is also presented. Finally, a class of T-user error-correcting codes for the noisy adder channel is constructed.

ON A CLASS OF δ-DECODABLE CODES FOR A MULTIPLE-ACCESS CHANNEL, Shu Lin (Department of Electrical Engineering, University of Hawaii, Honolulu, Hawaii 96822), Tadao Kasami (Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560, Japan), and Saburo Yamamura (Department of Measurement, Kobe University of Mercantile Marine, Kobe, Japan). Consider a multiple-access communication system with two users which employs two binary block codes of the same length n, C_1 and C_2. During a message interval, each user chooses a code word from his code and both users transmit simultaneously over a common channel. The channel combines the two transmitted code words into a single vector r over a certain alphabet. A single decoder at the receiving end processes r and decodes it in two code words, one in C_1 and the other in C_2. The code pair (C_1, C_2) used in the above system is referred to as a two-user code.

In this paper, we investigate block coding for a particular memoryless two-input single-output multiple-access channel. First, we present a class of t-error-correcting two-user codes which we call δ-decodable codes where $t = \lfloor(\delta-1)/2\rfloor$. Then, we derive lower bounds on the achievable rates of codes in this class for different ranges of t/n. We show that for certain range of t/n, there exist good 2-user codes (C_1,C_2) with rates (R_1,R_2) lying above the time-sharing line, i.e., $R_1+R_2 > 1$.

DETERMINISTIC CODES FOR SYNCHRONOUS AND ASYNCHRONOUS COMMUNICATION OVER THE REAL ADDER MULTIPLE ACCESS CHANNEL, Michael A. Deaett (General Electric Company, Pittsfield, Massachusetts) and Jack Keil Wolf (Department of Electrical Engineering, University of Massachusetts, Amherst, Massachusetts). Deterministic codes have been found for the multiple access real adder channel with L users. The cases of synchronized and of nonsynchronized inputs to the channel both are considered. Both variable length and block coding schemes are utilized. The rates of these codes are compared to the maximum achievable rates predicted by random coding.

SESSION A5

ON MEAN-SQUARE DECISION FEEDBACK EQUALIZATION AND TIMING PHASE, Jack Salz (Bell Telephone Laboratories, Holmdel, NJ 07733). We analyze the effect of timing phase on the performance of digital data receivers which employ decision feedback equalization. It has long been conjectured, and verified by computer simulations, that decision feedback equalizers are considerably less sensitive to the choice of timing phase than are conventional transversal linear equalizers. We develop the theoretical machinery which provides a rationale for these observations. It follows from our results that for typical operating conditions a 1-2 dB penalty can be incurred by choosing a bad timing phase in decision feedback, while the penalty can be an order of magnitude greater than this (in dB) in conventional linear equalizers.

PERFORMANCE OF AN ADAPTIVE EQUALIZATION TECHNIQUE FOR LINEAR FM SIGNALS, Laurence B. Milstein (Department of Applied Physics and Information Science, University of California at San Diego, La Jolla, California 92093). Linear FM signals have a variety of applications in communications and radar, most notably in pulse compression radars. However, they suffer serious distortion problems when used over channels which exhibit large phase nonlinearities. This is because the channel disperses the pulse in time over possibly many times its original width. Consequently, the need for equalization is clear, and adaptive equalization is desirable if one assumes the channel is slowly changing with time. In this paper, a new equalization technique designed for linear FM signals operating over channels characterized primarily by nonlinear phase characteristics will be described and analyzed (i.e., it will be assumed that the amount of phase distortion of the channel is significantly greater than the amount of amplitude distortion). The equalizer is designed to approximate as closely as possible the inverse of the channel so that the combination of it and a filter matched to the original transmitted waveshape will yield the desired pulse compression. For high input signal-to-noise ratios, a straightforward implementation of the above equalizer is presented.

APPLICATION OF FAST KALMAN ESTIMATION TO ADAPTIVE EQUALIZATION, D.D. Falconer (Bell Telephone Laboratories, Holmdel, NJ 07733) and L. Ljung (Department of Electrical Engineering, Linköping University, S-581 83 Linköping, Sweden). Very rapid initial convergence of the equalizer tap coefficients is a requirement of many data communication systems which employ adaptive equalizers to minimize intersymbol interference. As shown in recent papers by Godard, and by Gitlin and Magee, the Kalman estimation algorithm is applicable to the estimation of the optimal (minimum MSE) set of tap coefficients. It was furthermore shown to yield much faster equalizer convergence

than that achieved by the simple estimated gradient algorithm, especially for severely distored channels. We show how certain "fast Kalman estimation" techniques, originally introduced by Morf and Ljung, and with complexity proportional to the number of equalizer taps, can be adapted to the equalizer adjustment problem. The fast algorithm, applicable to both linear and decision feedback equalizers, exploits a certain shift-in-variance property of successive equalizer contents. It achieves the same fast convergence as the traditionally-implemented Kalman algorithm, whose complexity is proportional to the square of the number of taps. The rapid convergence properties of the fast Kalman adaptation algorithm are confirmed by simulation.

REDUCED-COMPLEXITY VITERBI ALGORITHM RECEIVER, Ghassan Kawas-Kaleh (Ecole Nationale Superieure des Telecommunications, 75634 Paris CEDEX 13, France). A generalized class of decoders, which contains both decision-feedback equalizer and maximum-likelihood Viterbi algorithm detector, is presented. A reduced complexity equalizer can be obtained by making the Viterbi algorithms work on a small number of the channel impulse response samples; the intersymbol interference resulting from other samples is estimated using tentative decisions available in the algorithm survivors.

The proposed equalizer has the advantage of effecting a decrease in residual intersymbol interference at the decoder input, when inserting a spectrum shaping prefilter, without increasing complexity.

Equalizer performances are analyzed, and simulation results are presented. They show a net superiority of the proposed equalizer compared to others.

A CLASS OF METHODS FOR COMPUTING ADAPATIVE EQUALIZER PARAMETERS, P. Cordes (FG Theorie der Signale, Technische Hochschule, Darmstadt, F.R., Germany). Transmitting digital data over bandlimited channels at high speed leads to distortions. In order to receive the data without error adaptive equalizers are used to equalize the channel.

Based on the mean-square error criterion methods for determining the parameters of adaptive equalizers have been specified and analyzed recently.

In this paper it is shown how most of these procedures can be combined to a common class. Systematic analysis of the optimal algorithm given by Kalman leads to new methods. They differ in computing effort and speed of convergence.

Computer simulations illustrates the results.

SESSION A5

THE EFFECTS OF LARGE INTERFERENCE ON DIGITALLY IMPLEMENTED ADAPTIVE ECHO CANCELLERS, R.D. Gitlin and S.B. Weinstein (Bell Telephone Laboratories, Holmdel, NJ 07733). Adaptive mean-square-tapped-delay-line echo cancellers, in either voice or data applications, are conventionally designed to stop adjustment during periods of "double-talking", i.e., when a large information-bearing signal is added to the echo to be cancelled. Continuous adjustment is, however, desirable in full-duplex, two-wire data transmission. We address the problem of tap adjustment, via the estimated-gradient algorithm, in the possible presence of a "double-talking" signal.

For an idealized double-talking model, it is demonstrated that the memoryless maximum-likelihood estimate of the residual echo is a linear function of the canceller output signal up to a certain threshold, beyond which the echo estimate falls back within a hysteresis region before evolving into another linear function with much smaller slope. This is nearly equivalent to the strategy of abruptly reducing the step size of the adjustment algorithm when double-talking begins, and is, in fact, an automatic mechanism for recognizing double-talking.

SESSION B1

ON DIGITAL SIGNATURES AND PUBLIC-KEY CRYPTOSYSTEMS, R.L. Rivest, A. Shamir and L. Adleman (MIT Laboratory for Computer Science). The operation of raising a number to a fixed power modulo a composite modulus is shown to be sufficient to implement "digital signatures": a way of creating for a (digitized) document a recognizable, unforgeable, document-dependent, digitized signature whose authenticity the signer can not later deny. This scheme has obvious applications in the design of "electronic funds transfer" systems or "electronic mail" systems, since here the messages must be digitized in order to be transmitted. Our approach is to provide an implementation of a "public-key crytosystem", an elegant concept invented by Diffie and Hellman. Such a system also enables enciphered communication between arbitrary pairs of people, without the necessity of agreeing on an enciphering key beforehand.

HIDING INFORMATION AND RECEIPTS IN TRAP DOOR KNAPSACKS, Ralph C. Merkle and Martin E. Hellman (Department of Electrical Engineering, Stanford University). Given a vector or integers $\underline{a}$ and a number S such that $S=\underline{a}\cdot\underline{x}$, where $\underline{x}$ in a binary vector, the knapsack problem is to find $\underline{x}$. Because this is an NP-complete problem it is strongly believed to be difficult to solve in general. A trap door knapsack vector $\underline{a}$ makes the problem appear very difficult unless one possesses trap door information used in the design of $\underline{a}$. Because only the designer can find the solution, others can send him information $\underline{x}$ hidden in the sum $S=\underline{a}\cdot\underline{x}$. This appraoch differs from usual cryptographic systems in that a secret key need not be exchanged.

Methods for generating and solving trap door knapsacks will be described. When implemented in circuitry these techniques allow anyone, regardless of mathematical ability, to automatically generate a trap door knapsack vector and to recover information sent to him in hidden form. A skilled opponent cannot discover this information even though he knows the general method used for generating the trap door vector $\underline{a}$.

It is also demonstrated that these techniques allow the generation of receipts (unforgeable, message dependent, digital signatures). Receipts are more valuable than usual digital authenticators because they allow the resolution of disputes between the transmitter and the receiver as to whether a particular message was sent.

AN IMPROVED ALGORITHM FOR COMPUTING LOGARITHMS OVER GF(p^m) AND ITS CRYPTOGRAPHIC SIGNIFICANCE, Stephen C. Pohlig and Martin E. Hellman (Department of Electrical Engineering, Stanford University, Stanford, CA 94305). An improved algorithm is presented for computing logarithms over GF(p), where p is a prime number. Previously published algorithms require $O(p^{1/2})$ complexity in both time and space.

The algorithm presented here requires $O(\sum_i n_i \cdot p_i^{1-r_i} \cdot \text{Log } p_i^{r_i})$ operations, $O(\sum_i p_i^{r_i} \cdot \text{Log } p_i)$ bits of memory, and $O(\sum_i p_i^{r_i} \cdot \text{Log } p_i^{r_i})$ precomputation operations, where $0 \le r_i \le 1$, and $p - 1 = p_1^{n_1} \ldots p_k^{n_k}$ is the prime factorization of $p - 1$.

A cryptographic system and a user authentication system based on exponentiation mod p are described. The above mentioned algorithm dictates that p - 1 should be chosen to have a large prime factor so that Logatithms over GF(p) are difficult to compute.

This improved algorithm directly extends to all finite fields $GF(p^m)$. Of practical interest is the case $GF(2^m)$ when 2^m-1 is prime (2^m-1 is a Mersenne prime). Exponentiation is not difficult to implement for this case. Many large Mersenne primes are known.

A HIGH-GRADE DATA ENCRYPTION ALGORITHM, Bernard J. Pankowski (Computer Sciences Corporation, 6565 Arlington Boulevard, Falls Church, Virginia 22046). A high-grade algorithm for encrypting data for transmission or storage is described. It develops more key space than the NBS algorithm, and offers an efficient software alternative to hardware encryption for applications where the volume of data requiring protection does not justify the cost of acquiring and maintaining dedicated encryption hardware.

Two algorithm variations have been coded, one processing only printable EBCDIC characters, and one scrambling all data forms (characters, integers, and floating-point numbers). Both versions produce "computationally secure" key stream ciphers by linking two congruence-type random generators to form a single long sequence (62 key bits), which combined with cipher feedback (6 or 8 key bits), produces a cryptographic key stream that never repeats.

Under control of this key stream, a variation of the Vigenere square (in use since 1586) performs either encryption or decryption using a polyalphabetic substitution on characters (or bytes). The Vigenere square variant is equivalent to an enciphered (simple substitution) modulo-n half-adder whose permutations, 64! for the character version and 256! for the all data version, add 378 and 2040 effective key bits respectively to the two algorithm variations.

No method of "breaking" the algorithm is known short of exhausting the key space, an economically prohibitive undertaking.

SESSION B1

PARALLEL PROCESSABLE CRYPTOGRAPHIC METHODS WITH UNBOUNDED PRACTICAL SECURITY, Jerome Rothstein (Department of Computer and Information Science, The Ohio State University, Columbus, Ohio 43210). Cryptographic methods are described which, at reasonable cost to users, impose prohibitive work penalties on code-breakers. The class of cryptographic transofrmations is so vast that codes can be used once and discarded, both in communication and in data security applications. We use the groupoid string formlism (earlier shown to combine Turing universality with parallel computing capability) specialized to quasigroups for unique decodability. Coding for error detection and correction can be done compatibly and independently. It appears likely that crypto and reliability encoding can be systematically combined to decoy code-breakers into semanticizing their decoding of reliability procedures into message opacifiers. Identification, authorization, and handshaking procedures can be integrated into the cryptographic method (with total update capability) at incremental user cost comparable to separate cost of those procedures or less.

AN ALGORITHM FOR SOLVING SIMPLE SUBSTITUTION CRYPTOGRAMS, Lalit Bahl, IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598). Al algorithm for automatic decipherment of simple substitution ciphers is presented. Let $\underline{X}$ denote the cleartext message, and $\underline{Y}$ the encrypted message. $\underline{X}$ is treated as the output of a Markov source, whose parameters are estimated from letter n-gram frequency counts obtained from a large sample of English text. The encryption process $\underline{X} \rightarrow \underline{Y}$ is modeled as transmission through a lossless channel with unknown transition probabilities $p(.|.)$. The transition matrix of this channel is, of course, a permutation matrix containing a single 1 in each row and column. Our aim is to determine $p(.|.)$, given $\underline{Y}$ and the source statistics.

Initially, the permutation nature of the channel transition matrix is ignored, and initial estimates of $p(.|.)$ are obtained from first order letter counts of $\underline{Y}$. The Forwards Backwards algorithm is used to iteratively refine these estimates. The algorithm adjusts $p(.|.)$ to increase $Pr\{\underline{Y}\}$ (which is a function of $p(.|.)$ and the source statistics). As entries in the channel transition matrix come close to 1, they are set to 1. The process stops when the transition matrix becomes a permutation matrix. This permutation is the key used for encrypting and its inverse can be applied to $\underline{Y}$ to obtain $\underline{X}$.

Using a letter 2-gram model of English, this algorithm conistently yielded the correct answer if the cryptogram was about 300 letters or longer. This, of course, is considerably longer than the 30 or so letters generally needed by humans to break simple substitution ciphers. Experiments with higher order n-gram Markov models are planned.

ENCRYPTION WITH RANDOMLY CHOSEN UNKNOWN FUNCTIONS, Ingemar Ingemarsson (Linköping University, Sweden). An encryption unit may be regarded as an unknown function. That is a set of functions $\{f_i(x)\}_{i=1}^{M}$ from a finite set X onto a finite set Y. The set is known to an outside observer. The key, i, and thus the actual function $f_i(x)$, however, is known solely by the legitimate user. It is assumed that the outside observer can test the encryption unit by applying r different inputs $x_1,\ldots,x_r$ and observing the corresponding outputs $y_1,\ldots,y_r$. Alternatively the inputs $x_1,\ldots,x_r$ are chosen randomly. Shannon studied the situation where the outside observer did not know the input ("the clear text") $x_1,\ldots,x_r$. In many cases it is more realistic to assume knowledge of a limited set of inputs. The question is to what extent this facilitates the cryptanalysis, i.e., the derivation of x from y. The key, i, is regarded as a random variable. We may then calculate the conditional probabilities $p[y|x,R]$ where R denotes the set of observed inputs and outputs $x_1,\ldots,x_r$ and $y_1,\ldots,y_r$. The ideal choice of unknown function from the designer's point of view, seems to be one where the output y is independent of R. Such unknown functions were called unknown functions with maximal uncertainty and analyzed by this author.

The aim of this paper is to study unknown functions where the M functions $f_i(x)$ are chosen randomly from the set of all possible functions from the finite set X onto a finite set Y. The main result is that there is a remarkable and very fast convergence with increasing M in that a vast majority of all unknown functions belong to a class with nearly the same conditional probabilities $p[y|x,R]$. Thus even if it were shown that there exist unknown functions with maximal uncertainty, nearly all unknown functions have less than maximam uncertainty. For small r, however, the uncertainty is close to maximal. (Uncertainty is here measured as the conditional entropy.) Randomization is not only a mathematical tool. It can also be used to produce unknown functions, to be used, for example, in identity control units. The convergence asserts that almost all units produced in such a way have similar statistical properties.

When unknown functions are used for encryption we must require that the functions $f_i(x)$ are invertible. Hence the analysis described above is extended to invertible unknown functions, i.e., unknown functions where:

$$f_i(x_m) \neq f_i(x_n) \text{ for all i and } x_m \neq x_n$$

Finally results are also given for unknown functions which are randomly chosen without replacement from the set of all possible functions. This guarantees that all M functions in the set $\{f_i(x)\}_{i=1}^{M}$, are distinct.

BOUNDS ON SECRECY SYSTEMS, Shyue-Ching Lu (Telecommunication Laboratories, Ministry of Communications, R.O.C.). The problem of enciphering the output of a discrete message source is considered. The key used in enciphering is transmitted safely to the user; however, the enciphered message is intercepted by an "enemy".

The class of additive-like instantaneous block (ALIB) encipherers is studies. An ALIB encipherer is specified by an additive-like combiner and a set of key words. An ALIB encipher operates as follows: the selected key determines a particular key word and the cryptogram is produced by the combiner letter by letter from the message and the particular key word. The strength of an encipherer is measured by the probability of correct decryptment, the probability that the non-zero per letter Hamming distortion is less than the tolerable distortion, the number of keys and the tolerable distortion. It is shown that there exists an ALIB encipherer with key rate r, tolerable distortion Δ such that both the probability of correct decryptment, p_w, and the probability that the non-zero per letter Hamming distortion is less than the tolerable distortion; $p(\Delta)$, can be made arbitrarily small provided the key rate is greater than a bound on the cipher-distortion function.

The concept of sphere packing is used to derive both upper and lower bounds on p_w for binary additive instantaneous block encipherers and memoryless binary sources. In this case $p(\Delta)$ is required to be zero for a prescribed tolerable per letter Hamming distortion Δ. For ALIB encipherers and discrete message sources the concepts of random ciphering and exponential bounding are employed to obtain bounds on p_w and $p(\Delta)$.

In contrast to Shannon"s "random" cipher result, it is shown that good encipherers exist with key rates less than the message redundancy.

ON A NEW CLASS OF OPTIMAL DETECTORS: THE APPROXIMATE IMPULSE DETECTION, Michael Negin and C. Britton Rorabaugh, Jr. (Department of Electrical Engineering, Drexel University, 32nd and Chestnut Streets, Philadelphia, PA 19104). Matched filters are often used for detecting a signal of known shape, but with a random time of arrival. In the case where multiple signals are to be detected, the matched filter is useful as long as the multiple signals do not overlap or superimpose. Since the matched filter is optimized with respect to an amplitude criterion (peak output signal to root mean squared output noise), it is no surprise that the matched filter can perform poorly when signals overlap in time. Several investigators have proposed modified inverse filters to solve the problem of overlapped signals. Since in a noise-free environment, inverse filters produce impulses for every occurrence of the signal to be detected, it seems reasonable to try to use them as detectors. Unfortunately in the noisy environment, inverse filters usually amplify noise considerably. The modified inverse filters discussed in literature have typically not been general in development and furthermore usually require information such as a desired resolution time.

The approach taken in this paper is one that formulates an optimal detection criterion that combines amplitude and resolution criteria.

The criterion essentially results in a linear filter that (in the mean square sense) optimally estimates an impulse train that corresponds to the time of arrival statistics of the signal train to be detected. Multiple and overlapping signals are optimally detected since the time-of-arrival statistics are used in deriving the optimal filter.

The Approximate Impulse Detector (AID) is given by,

$$F_o(\omega) = \frac{\phi_{pp}(\omega)\ S^*(\omega)}{\phi_{pp}(\omega)\ \phi_{ss}(\omega) + \phi_{nn}(\omega)}$$

where $F_o(\omega)$ is the optimal filter, $S^*(\omega)$ is the conjugate of the signal to be detected, $\phi_{nn}(\omega)$ is the power spectral density of the noise, and $\phi_{ss}(\omega) = S(\omega)\ S^*(\omega)$. $\phi_{pp}(\omega)$ is the power spectral density of an impulse train that corresponds to the time of arrival statistics of the signal train to be detected.

The optimal filter has been simulated in a variety of signal and noise conditions and indeed has considerably better resolution properties than the matched filter. The signal to noise ratio performance of the AID filter is comparable to the matched filter. Synthesis of AID filters is only slightly more difficult than matched filter synthesis. Furthermore it can be seen from the equation that in certain high noise environments, ($\phi_{nn} \to \infty$) AID can reduce to the

matched filter. In low noise environments ($\phi_{nn} \to 0$), AID yields an inverse filter. In moderate noise environments, the AID filter specifies an optimal detection filter for a given noise and signal environment.

SEQUENTIAL PARTITION DETECTORS WITH DEPENDENT SAMPLING, Roger F. Dwyer (Naval Underwater Systems Center, New London, CT 06320) and Ludwik Kurz (Polytechnic Institute of New York, Department of Engineering and Electrophysics, Brooklyn, NY 11201). In this paper, the theory of sequential partition detectors with dependent sampling is introduced. A new formulation is given which predicts the thresholds under q-dependent sampling in order to maintain the same error probabilities as in the independent sampling case. A comparison is made between independent and dependent sequential partition detectors based on the average time to detection. Under stated conditions dependent sequential partition detectors show improved efficiency for both Lehmann and shift of the mean alternatives.

A NEW APPROACH TO ADAPTIVE DETECTION, B. Picinbono (Laboratoire des Signaux et Systèmes, ESE - Plateau du Moulon, 91190, GIF, France). The statistical detection theory of signal in noise is generally presented in a non adaptive context. In such a context is is assumed that some statistical properties of the signals and the noise are known beforehand. For example the likelihood optimal receiver needs the knowledge of the probability distributions. Unfortunately it is often difficult to have such a knowledge, and moreover there are in general important variations of the statistical parameters which make an adaptive processing necessary.

It is difficult to present a general theory of adaptive and optimal detection. The most common way, but not necessarily optimal, is to replace in the structure of optimal receivers the unknown parameter by a convenient estimation. But to perform this estimation we need a sample of the observation which is composed of only noise. Another way is to extract from the observation a function which is a Noise Alone Reference (N A R) and which can give informations on the noise, even if a signal is present.

In this paper we introduct the notion of N A R and we give some examples appearing in detection problems. Moreover we show that the notion of N A R appears very naturally in the physical interpretation of matched filter in colored noise. Finally we give some indications of the use of such a N A R in detection of deterministic signals in noise with fluctuating power.

SESSION B2

STOCHASTIC APPROXIMATION IN DETECTION, S.C. Lee and L.W. Nolte (Department of Electrical Engineering, Duke University, Durham, NC 27706). The role of stochastic approximation in detection theory is investigated for the Gauss-Gauss array processing problem. The adaptive least mean-square (LMS) algorithm is used as a stochastic approximation technique for both detection and estimation. In the Gauss-Gauss problem, it is well known that the optimum detector can be put into the form of an estimator-correlator in which the minimum mean-square error estimator (MMSE) is correlated with the data. A suboptimum detector is considered here in which the adaptive LMS algorithm is used as the estimator in an estimator-correlator structure.

Detection performance is analyzed using the detectability index both for optimum and suboptimum detectors. The estimation performance of the MMSE, MLE, and adaptive LMS estimators are also computed using the mean-square error. It is shown that the detectability of the suboptimum detector suffers seriously compared to the detectability of the optimum detector. The mean-square error of the adaptive LMS estimator, however, converged to that of the optimum estimator (MMSE) within a reasonable number of iterations under the signal-present hypothesis.

ROBUST SEQUENTIAL DETECTION OF SIGNALS IN NOISE, A.H. El-Sawy and V.D. VandeLinde (Department of Electrical Engineering, The John Hopkins University, Baltimore, Maryland 21218). The problem of sequential detection of weak signals in additive noise is solved under the assumption that the unknown noise density function is a member of some known class of symmetric densities. Two general approaches to the design of receivers which as asymptotically most robust in a minimax sense are established. Both receivers guarantee an upper bound on the probability of errors of the first and second type and minimize the maximum expected risk in Bayes sense. The main difference between them is that the first, which we call the M-sequential detector, requires a growing memory, and the second, which is called the stochastic approximation sequential detector, does not. A comparison between the two methods is provided for the special case when the noise density family is defined by

$$F = \left(f: \int_{-a}^{a} f(x)\,dx = p,\ f \text{ symmetric and continuous at } \pm a\right)$$

In addition the savings in the expected sample size over some non-sequential methods have been calculated.

SESSION B2

SIGNAL DETECTION THROUGH NOISE REMOVAL, A.J. Claus, T.T. Kadota, and D.M. Romain (Bell Telephone Laboratories, Inc., Murray Hill and Whippany, NJ). This paper deals with the construction of an orthonormal set of vectors which is useful for adaptive side lobe manipulation of array beam patterns. The vectors allow side lobe suppression across any spatial sector and they depend only on the array geometry. The latter feature permits the numerical values of vector components to be pre-computed and permanently stored making their use attractive for real-time adaptive array processing. The center of the sector across which side lobes must be altered determines array sensor steering delays, and the sector width specifies the number of vectors to be used.

This number is closely related to the essential rank of some frequency dependent matrix. The matrix is subjected to a sequence of linear mappings, each mapping producing a vector of the orthonormal set. The mappings are carried out at frequencies determined by preassigned error bounds. This procedure results in frequency independent vectors capable of side lobe manipulation to within any degree of accuracy.

BOUNDS ON PROBABILITY OF ERROR AND SUBOPTIMAL DETECTORS, Alain Fogel and Stuart C. Schwartz (Department of Electrical Engineering and Computer Science, Princeton University, Princeton, NJ 08540). A wide class of optimum likelihood ratio detectors are specified in terms of a conditional mean estimate (CME), and the resulting implementation is an estimator-correlator. When the CME is either unknown or difficult to evaluate, other estimates can be substituted, resulting in a class of sub-optimum detectors. Although probabilities of error are the central measure of performance of a detector, these are usually not available and bounding techniques such as the Chernoff bounds have provided an alternative approach to obtain a tractable criterion of performance. Using a variational approach, we investigate the relationship between the two resulting Chernoff bounds on the probabilities of error for the above class of optimum and suboptimum detectors. Problems considered include: continuous-time detection of a signal in white Gaussian noise; signals modulating the rate of a doubly stochastic Poisson process; the discrete-time estimator correlator known to be optimum when the noise distribution belongs to the exponential family.

SESSION B3

SIGNAL FILTERING WITH GAUSSIAN QUANTUM STATES AND CANONICAL MEASUREMENT, R.O. Harger and J.S. Baras (Electrical Engineering Department, University of Maryland, College Park, Maryland 20742). The recently initiated extension of classical signal filtering accounting for quantum mechanical limitations on measurement is here continued incorporting measurement effect on a quantum state which depends on a vector random "signal" process and evolves according to the Shrödinger equation. The quantum measurements are restricted to be "canonical" and a system Hamiltonian that results in a Gauss-Markov measurement outcome sequence is assumed. The weighted mean-square error is minimized over choice of current quantum measurement and linear weightings of past measurements. It is shown that the optimal weight matrices satisfy certain normal equations and necessary and sufficient conditions for the optimal canonical measurement are given. Sufficient conditions for the "separation" into a canonical measurement independent of past measurements and classical "post" filtering are given. An optical communication application is given.

QUANTUM STATE PROPAGATION AND OPTICAL CHANNELS, Horace P. Yuen (Research Laboratory of Electronics, Cambridge, Massachusetts 02139) and Jeffrey H. Shapiro (Research Laboratory of Electronics and Department of Electrical Engineering and Computer Science, M.I.T., Cambridge, Massachusetts 02139). To determine the ultimate quantum limitations on the performance of optical communication systems, it is essential to consider optimum quantum state generation at the transmitter and the effect of the propagation channel on the received state. The propagation of arbitrary quantum states in free space is developed herein as a quantum diffraction theory. This theory yields the correct classical Huygens-Fresnel limit, and justifies the usual semiclassical propagation treatment of coherent state fields. For states that cannot be described by a well-behaved P-representation, such as the recently discussed two-photon coherent states, quantum diffraction theory permits separation of the state-generation problem from the state-propagation problem.

This allows the quantum transmitter, the channel, and the receiver to be represented in a convenient framework similar to the semiclassical theory of optical communication. This framework will be used to examine near-field and far-field performance potential of two-photon coherent states. Application of the same approach to some other optical channels will be indicated.

SESSION B3

STRUCTURED OPTICAL DETECTION OF TWO-PHOTON RADIATION, Jesus A. Machado Mata, Jeffrey H. Shapiro and Horace P. Yuen (Department of Electrical Engineering and Computer Science and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139). Recent theoretical work has shown that novel quantum states, called two-photon coherent states (TCS), have significant potential for improving free-space optical communications. Because TCS radiation does not possess a classical analog, the usual semiclassical statistical models for photodetection are not applicable to TCS reception. In this paper, we present performance results for direct, heterodyne, and homodyne detection of TCS radiation. We shall show that TCS radiation offers a modest energy advantage (at most 3 dB) over coherent-state light in direct detection of binary pulse-position modulated digital signals, and in heterodyne detection of amplitude modulated analog signals. A much more significant improvement will be exhibited for the case of homodyne detected amplitude modulated signals. In particular, under near-field propagation conditions TCS homodyne signal-to-noise ratio will exceed that for coherent-state homodyning by a factor of $N_S + 1$, where N_S is the average number of signal photons. Under far-field propagation conditions, assuming proper mode-mixing with a TCS local oscillator in the vicinity of the receiver, homodyne detection can attenuate post-measurement quantum noise by a factor equal to the radiative loss encountered in the channel. The effect of finite quantum efficiency on these results will be discussed.

MAXIMUM LIKELIHOOD SEQUENCE ESTIMATION FOR RANDOMLY DISPERSIVE OPTI-OPTICAL-COMMUNICATION CHANNELS, Robert E. Morley, Jr. (Micro-Term, Inc., St. Louis, Missouri 63117) and Donald L. Synder (Department of Electrical Engineering and Biomedical Computer Laboratory, Washington University, St. Louis, Missouri 63130). Receiver designs for maximum likihood sequence estimation of digital data transmission through randomly dispersive optical-communication channels are developed. The random channels are constrained to have finite memory in the sense that the casual minimum mean square error estimate of the channel output at any given time is a function of only a finite amount of past data and a finite number of information symbols. Examples of the channels included are the time varying Rayleigh, Ricean and lognormal fading channels whose covariance functions have finite support and phase unsynchronized channels which satisfy the finite-memory constraint. The maximum likelihood sequence estimation receivers developed are optimal in the sense that they minimize the probability of selecting the incorrect message. Optimal receivers are presented not only for observations modeled by continuous processes, such as encountered with hetrodyne receivers, but also for observations modeled by point processes, such as those encountered with direct-detection receivers. The information sequence may be either a stream of independent, equally-likely

symbols of binary or M-ary alphabets or the output sequence of a trellis or convolutional encoder. An upper bound on the average message error-probability is derived.

QUANTUM OPTIMIZED RANDOM CODING EXPONENT FOR (N,R) BLOCK-CODED BINARY PURE STATES, Nam-Soo Myung (Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139). The random coding bound for binary (N,R) block-coded sequence of pure quantum states is considered. Lagrange multiplier method is employed to determine the optimum detection operator at each band. It is found that the measurement which minimizes the random coding baud is the same as the optimum measurement for detecting equally probable pure state binary signals. Comparison of random coding exponent among different receivers is made under signal energy and band width constraints.

MAXIMIZATION OF MUTUAL INFORMATION AND MINIMIZATION OF DETECTION ERROR IN QUANTUM COMMUNICATIONS, Vincent W.S. Chan (Department of Electrical Engineering, Cornell University, Ithaca, New York 14853). In this paper, we will consider the maximization of the mutual information (I(g;J)) between the set of possible quantum system states (g) and the set of possible outcomes (J) of a quantum measurement. The maximization is over the choice of all feasible quantum measurements (i.e., measurements characterized by operator-valued measures). Explicit solutions have been found for the binary pure state problem and the M-ary problem with pairwise commuting density operators. In both cases the maximizing measurement is a 'complete' measurement (i.e., the system is left in a known eigenstate after measurement). Hence the number of possible outcomes N can be larger than M, the number of possible states. In fact N can be as large as $M \cdot \dim\{H\}$, where H is the Hilbert Space spanned by the states of the system. We found, in both cases, the measurement that maximizes mutual information also minimizes detection error probabilities, although the converse is only true for the binary pure state problem. For the M-ary problem with pairwise commuting density operators the measurement that maximizes I(g;J) has $N = \dim\{H\}$ possible outcomes, with N possibly larger or less than M. It can be shown that the measurement that minimizes $Pr[\varepsilon]$ may not maximize I(g;J).

A FAST EVALUATION OF ERROR RATE IN FIBER OPTIC DIGITAL SYSTEMS WITH CORRELATED SYMBOLS, Nevio Benvenuto and Silvano G. Pupolin (University of Padova, Department of Electrical Engineering, Via Gradenigo 6A, 35100 - PADOVA, Italy). A fast evaluation of the error rate in fiber optic digital transmission systems with mutually independent

symbols has been proposed. The method is based on a Gram-Charlier series expansion and requires the joint moments of the digital signal and shot noise variance (this variance is linearly related to the data message).

The present paper extends the above method to systems with correlated symbols, where correlation is produced by line encoding. Using a sequential machine representation for line encoder a fast technique of moment evaluation is set up, where the computational work increases linearly with the number of interferers, thus avoiding the exponential growth of exhaustive method.

SESSION B4

TRELLIS CODING FOR DEGRADED BROADCAST CHANNELS, W.J. Leighton, III and Harry H. Tan (Department of Electrical Engineering and Computer Science, Princeton University). Coding theorems are proved for two-receiver degraded discrete memoryless broadcast channels that use superposition trellis encoders and (1) Viterbi-algorithm type decoders at both receivers or (2) Stack sequential type decoders at both receivers. In the latter case, it is also shown that the expected number of computations per decoded sublock is finite at all rates inside a rate region R_{COMP} contained in the capacity region. Finally, R_{COMP} is computed for a degraded binary symmetric broadcast channel and is shown to be properly contained in the capacity region.

AN ACHIEVABLE RATE REGION FOR THE MULTIPLE-ACCESS CHANNEL WITH FEEDBACK, Thomas M. Cover (Department of Electrical Engineering, Stanford University, Stanford, CA 94305) and S.K. Leung-Yan-Cheong (Department of Electrical Engineering, M.I.T., Cambridge, MA 02139). An achievable rate region $R_1 \leq I(X_1;Y|X_2,U)$, $R_2 \leq I(X_2;Y|X_1,U)$, $R_1 + R_2 \leq I(X_1,X_2;Y)$, is exhibited for the multiple-access channel with feedback. This region exceeds the achievable rate region without feedback and exceeds the rate point found by Gaarder and Wolf for the binary erasure multiple-access channel with feedback.

The presence of feedback allows the independent transmitters to understand each other's intended transmissions before the receiver has sufficient information to achieve the desired decoding. This allows the transmitters to cooperate in the transmission of information resolving the residual uncertainty of the receiver. At the same time, fresh independent information from the transmitters is superimposed on the cooperative correction information.

FEEDBACK CAPACITY OF DEGRADED BROADCAST CHANNELS, Abbas El Gamal (Information Systems Laboratory, Stanford University, Stanford, CA 94305). We consider the model of one sender and two receivers connected through two cascaded discrete memoryless channels $\{P_1(Y|X)\}$ and $\{P_2(Z|Y)\}$. We prove that the addition of casual noiseless feedback from both receivers to the sender does not enlarge the capacity region of the channel. The result is intuitively clear since feedback merely adds degraded forms of the past sender sequence. It is also consistent with Shannon's result on the discrete memoryless channel with feedback.

We next generalize Shannon's entropy power inequality for the additive white Gaussian noise channel to the case when casual noiseless feedback is added. Then, using Hadamard's inequality we derive an upper bound on the entropy of the output sequence of the AWGN channel with feedback. The derived bounds are used to prove that the capacity of the degraded AWGN broadcast channel is not increased by feedback.

ON THE COMPUTABILITY OF AN ACHIEVABLE RATE REGION FOR THE BINARY-INPUT BROADCAST CHANNEL, B.E. Hajek and M.B. Pursley (Coordinated Science Laboratory, University of Illinois, Urbana, Illinois 61801). Cover and van der Meulen independently established an achievable rate region, $\mathcal{R}$, for a general discrete memoryless broadcast channel that transmits separate messages at rates R_1 and R_2 to each of the two receivers and a common message at rate R_0 to both receivers.

The capacity region $\mathcal{R}^*$ satisfies the following internal consistency condition: If $(R_1,R_2,R_0) \in \mathcal{R}^*$, $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, and $\lambda_1 + \lambda_2 \leq 1$ then

$$(R_1 + \lambda_1 R_0,\ R_2 + \lambda_2 R_0,\ (1 - \lambda_1 - \lambda_2)R_0) \in \mathcal{R}^*.$$

Our results imply that $\mathcal{R}$ does not satisfy this condition and can therefore be enlarged. Let $\hat{\mathcal{R}}$ denote the smallest region which contains $\mathcal{R}$ and is internally consistent in the above sense.

The inconsistency of $\mathcal{R}$ is deduced from investigation of $\mathcal{R}_0$ and $\hat{\mathcal{R}}_0$, the projections onto the $R_0 = 0$ plane of $\mathcal{R}$ and $\hat{\mathcal{R}}$, respectively. These correspond to the situation in which there is no common message.

We show that $\hat{\mathcal{R}}_0$ is the closed convex hull of the set of all (R_1,R_2) such that

$$R_i \leq I(U_i,U_0;Y_i) \text{ for } i = 1,2$$

and

$$R_1 + R_2 \leq \sum_{i=1}^{2} I(U_i;Y_i|U_0) + \min\{I(U_0;Y_j)|j=1,2\}$$

for some Markov chain $(\underline{U},X,\underline{Y})$ such that $\underline{U} = (U_1,U_2,U_0)$ are mutually independent and $\underline{Y} = (Y_1,Y_2)$ is a broadcast channel output corresponding to input X. For binary-input channels, we show that $\hat{\mathcal{R}}_0$ can be calculated by considering only those $(\underline{U},X,\underline{Y})$ for which U_0 is binary, U_1 and U_2 are ternary, and X is a (deterministic) function of $\underline{U}$, and that $\mathcal{R}_0$ can be calculated by considering only binary auxiliary random variables. We present examples for which $\hat{\mathcal{R}}_0$ is strictly larger than $\mathcal{R}_0$ and hence $\hat{\mathcal{R}}$ is strictly larger than $\mathcal{R}$.

SESSION B4

THE GAUSSIAN WIRETAP CHANNEL, S.K. Leung-Yan-Cheong (Electronic Systems Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139) and Martin E. Hellman (Department of Electrical Engineering, Stanford University, Stanford, CA 94305). In a recent paper, Wyner introduced the concept of the wiretap channel. He considered the case in which data is to be transmitted reliably over a discrete memoryless channel (referred to as the main channel) to a legitimate receiver. The wiretapper views the output of the main channel through another discrete memoryless channel. It is assumed that the wiretapper knows the encoding scheme used at the transmitter and the decoding scheme used by the legitimate receiver. The object is to design the encoder-decoder so as to maximize the transmission rate R to the legitimate receiver and the equivocation d of the data at the wiretapper.

In this note, Wyner's results are extended to the Gaussian wiretap channel. The R vs. d curve is determined explicitly through the use of some special properties of the Gaussian channel. It is shown that the complete set of achievable (R,d) pairs is given by $\{(R,d) \mid R \leq C_M,\ d \leq 1,\ Rd \leq C_S\}$ where C_M is the capacity of the main channel and C_S is the difference between the capacities of the main and wiretap channels.

SESSION B5

THE SOURCE CODING THEOREM REVISITED: A COMBINATORIAL APPROACH, Giuseppe Longo (Instituto di Elettrotecnica ed Elettronica, Università di Trieste, Trieste, and Centre International des Sciences Mécaniques, Udine, Italy) and Andrea Sgarro (Instituto de Elettrotecnica ed Elettronica, Università di Trieste, Trieste, Italy). In this paper a combinatorial approach is proposed to the classical source coding problems for a finite memoryless stationary source (achievable rates and error probability exponent). This approach provides a sound heuristical justification for the widespread appearance of entropy and divergence (Kullback's discrimination) in source coding. Use is made of the fact that asymptotically the rate of any composition class is precisely an entropy, while its normalized probability exponent is precisely a divergence; similar results hold also for any union of composition classes. (A composition class is a set made up by all the source sequences which are permutations of one another.) These remarks allow us to derive the classical results by "approximating" the relevant sets by means of unions of composition classes. No use is made either of the law of large numbers or of Chebyshev inequality.

SOURCE CODING TO MINIMIZE THE PROBABILITY OF BUFFER OVERFLOW, Pierre A. Humblet (Electronic Systems Laboratory, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139). A discrete memoryless asychronous source emits symbols that are encoded one by one and stored in a buffer before being transmitted over a synchronous output line. For each source code, the probability that the arrival of a symbol causes a buffer overflow decreases roughly exponentially with the size of the buffer. If symbol i has probability p_i and is encoded into a codeword of length m_i, and if the Laplace-Stieltjes transform of intermission time distribution is A(s), the rate of exponential decrease is related to the supremum of the s such that $A(s) \sum_i p_i e^{sm_i} \leq 1$.

We develop an iterative algorithm for constructing a prefix condition code that maximizes this supremum. The algorithm has two main parts. In each iteration the first part finds a prefix condition code minimizing $\sum_i p_i e^{sm_i}$ for a $s \geq 0$ given by the second part during the previous iteration; it is a generalization of Huffman's encoding procedure. The second part computes the supremum defined above for the code generated by the first part. We show that this algorithm yields an optimal code in a finite time. We also give an upper bound on the maximum of the supremum.

SESSION B5

ON VARIABLE-TO-VARIABLE CODING FOR DISCRETE MEMORYLESS SOURCES, Leonardus Th. N.M. Mulder and David L. Cohn (Electrical Engineering Department, University of Notre Dame, Notre Dame , Indiana 46556). The objective of any noiseless source-coding algorithm is to minimize the rate of the coding-scheme at some fixed level of complexity. We consider a fixed number of code words n. Optimal variable-to-block and block-to-variable schemes are known. As can be shown, a combination of both schemes does not lead to an optimal variable-to-variable scheme.

For $n \leq 10$ optimal codes for binary sources were found by enumeration of different-weighted binary trees. For $n > 10$ and for non-binary sources, enumeration becomes prohibitive. As an alternative, sub-optimal algorithms for finding good codes have been developed. In these algorithms, 'worst'-nodes in the encoding tree are extended. A criterion for 'worst'-node is derived. The algorithms are computationally efficient and the resulting codes out-perform the combination of optimal variable-to-block and block-to-variable schemes for most sources. For low entropy sources, the combination scheme becomes run-length encoding which turns out to be optimal.

BOUNDS ON THE COST OF OPTIMAL UNIQUELY DECIPHERABLE CODES, Norbert Cot and John Gill (Electrical Engineering Department, Stanford University, Stanford, CA 94305). Let $\{b_1,\ldots,b_t\}$ be the positive costs of the symbols of a t-ary alphabet. By the converse to Shannon's noiseless coding theorem, a lower bound on the average codeword cost for any uniquely decipherable code with codeword probabilities is $H(p_1,\ldots,p_n)/\log_2\lambda$, where $H(p_1,\ldots,p_n)$ is the entropy in bits of the codeword ensemble and λ is the unique positive root of the equation $z^{-b_1} + \ldots + z^{-b_t} = 1$. We generalize Shannon's well known technique for constructing nearly optimal prefix codes and obtain as an upper bound for the cost of optimal uniquely decipherable codes $H(p_1,\ldots,p_n)/\log_2\lambda + b_1 + b_2 \log_\lambda(\lambda_2) + b_3 \log_\lambda(\lambda_3/\lambda_2) + \ldots + b_t \log_\lambda(\lambda_t/\lambda_{t-1})$, where λ_j is the unique positive root of the equation $z^{-b_1} + \ldots + z^{-b_j} = 1$.

ADAPTIVE HUFFMAN CODES, Robert G. Gallager (Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science and Electronic Systems Laboratory). In honor of the twenty-fifth anniversary of Huffman Coding, three new results about Huffman codes are presented. The first result shows that a binary prefix condition code is a Huffman code (with a convention of assigning 0 to the more probable branch in the construction procedure) if the probabilities of the intermediate and terminal nodes in the code tree are

lexicographically ordered. The second result upper bounds the redundancy (expected length minus entropy) of a binary Huffman code by $P_1 + \log_2[2(\log_2 e)/e] \approx P_1 + .086$ where P_1 is the probability of the most likely source letter. The third result is a simple algorithm for adapting a Huffman code to slowly varying estimates of the source probabilities. In essence, one maintains a running count of uses of each node in the code tree. Whenever the occurrence of a message increases a node count above the count of the next lower node in the lexicographic ordering, the nodes, with their attached subtrees, are interchanged.

AN EQUAL-LENGTH-AT-OUTPUT UNIVERSAL METHOD OF CODING, B. Fitingof (Technion-Israel Institute of Technology, Haifa, Israel). Three different concepts of optimal coding method are defined and compared: The asymptotically optimal method of coding (AOMC), the asymptotically optimal on the average method of coding (AOAMC) and the universal method of coding for a class of sources (UMC). A sufficient condition of universality for the class of Bernoullian sources with a given alphabet is found in terms of quasi-entropy of input words. The notion of a monotone source is introduced and used in the construction of a called K-method, where input words of various lengths are encoded in output words of equal length. It is proved that the K-method is an AOMC for any monotone sources, in particular, for a Bernoullian source. A special case of the K-method is considered and proved to be a UMC for the class of all Bernoullian sources with a given size of alphabet. An algorithm implementing the K-method of coding is suggested, which requires very low storage.

PLENARY SESSION B

ALGEBRAIC COMPLEXITY OF COMPUTATION, S. Winograd (IBM Research, Yorktown Heights, NY). This talk will survey some recent results in the theory of algebraic complexity of computation with particular emphasis on problems connected with signal processing and coding.

SESSION C1

COMPUTATIONAL COMPLEXITY OF CONVOLUTIONS EVALUATED BY NUMBER THEORETIC TRANSFORMS, H. Nussbaumer (IBM CER, 06610 La Gaude, France). Various Number Theoretic Transforms (NTT) have been introduced recently as a means of reducing the computational complexity of convolutions and correlations. In this paper, we make precise the processing requirements for pseudo Mersenne and pseudo Fermat number transforms. These results are compared to those corresponding to Mersenne and Fermat number transforms and extended to cover the case of complex transforms.

The computational complexity of non recursive digital filters implemented with these various transforms is evaluated and compared, under certain simplifying assumptions, to that corresponding to direct computation. We show that the processing efficiency of the NTT approach increases with the required precision and that optimum results are obtained with Fermat number transforms and the complex pseudo Mersenne and pseudo Fermat number transforms defined modulo $\frac{(2^q-1)^2}{2^q-1}$ and modulo $\frac{(2^q+1)^2}{2^q-1}$ with q prime. The computational complexity as a function of filter length is investigated and it is shown that one-dimensional filters having a length of up to about 300 taps can be implemented efficiently with NTT's.

FAST ALGORITHMS FOR MULTI-VARIABLE AND MULTI-DIMENSIONAL SYSTEMS, M. Morf (Information Systems Laboratory, Stanford University, Stanford, CA 94305). The computational complexity and efficiency of algorithms for the analysis, design and signal processing involving single and multi-variable and multi-dimensional systems depends strongly on the structure of such systems. Most algorithms are based on matrix multiplications, factorizations or inversions. The structure of the underlying systems is reflected in the matrix representations of these operations. Often the structure of matrices is visible by inspection. For instance the Toeplitz or Hankel matrices reflect stationarity of signals or time-invariance of systems; the matrix entries are functions of the difference or the sum of the indices. However, in general, operations such as sums of products or inverses of matrices with an obvious structure often do

not show a particular pattern. If the structure can be exposed, very likely it can be used to reduce the computational complexity and number of operations, i.e., increase the efficiency of algorithms for matrix calculations.

A particular class of considerable interest to the signal processing and systems area are the matrices that can be interpreted as sub-matrices of the coefficient matrix of rational polynomials in two or more variables ("2 or M-D transforms"). Sums of products of Toeplitz or Hankel matrices represent one particular case, where the denominator of the associated "2-D transform" is just the difference of the two polynomial variables. A special case of this are the Christophel-Darboux formulas for the resolvent associated with a Toeplitz kernel. Another particular case of recent interest are the variable Fast Fourier and related transforms,e.g., the Chirp Z-Transform corresponds to a factorization of the FFT matrix into a product of Diagonal and Toeplitz matrices. A crucial property of these matrices is, however, that they can be factored into Kronecker products of simpler matrices; these products in turn correspond to (primitive) factors of the associated "2-D transform". We have now algorithms that can find rational approximations of functions in two variables efficiently, thus a method to expose the structure of a given matrix. The area of 2-D (matrix) polynomials has recently grown very rapidly and the newly available results can shed light on the connections between the various matrix results, the associated fast algorithms and their complexity. A survey of these results and, in particular, relations between displacement matrices (Toeplitz and Hankel) and FFT algorithms will be presented.

THE USE OF DECISION TREES IN COMPUTATIONAL COMPLEXITY, Andrew C. Yao (Department of Computer Science, Stanford University, Stanford, CA 94305). For a given computational problem, its complexity is understood through devising efficient algorithms to find solutions, on the one hand, and proving lower bounds to the number of steps required by all algorithms, on the other. In discussing lower bounds, one must first adopt a model suitable for representing all possible algorithms under consideration. The decision tree model has proven to be appropriate for a great variety of problems, such as information retrieval, string matching, finding shortest paths, etc., and many fruitful results obtained. In this talk, we present an overview of the decision tree model. Examples are discussed to illustrate the interesting techniques.

SESSION C1

FAST STATISTICAL ALGORITHMS, M. Shamos (Department of Computer Science, Carnegie-Mellon University, Pittsburgh, PA 15213). We will discuss techniques for designing efficient algorithms for statistical computation, taking examples from nonparametric estimation, correlation, regression and time series analysis. We will stress on-line methods and average-case analysis in developing a set of fast computational tools that can be used to build more sophisticated algorithms. Attendees will be encouraged to pose problems that they have encountered in applications.

SESSION C2

NONPARAMETRIC ESTIMATION WITH LOCAL RULES, C.S. Penrod and T.J. Wagner (Department of Electrical Engineering, University of Texas, Austin, Texas 78712). The application of nearest neighbor rules and other local rules to the problem of estimating a parameter θ is investigated. It is assumed that a loss function L, an observed random vector X, and data consisting of a sequence of independent random vectors $(X_1,\theta_1),\ldots,(X_n,\theta_n)$ with the same distribution as (X,θ) are given. Conditions are shown for which, if R* denotes the Bayes risk (the minimum expected loss possible), then the conditional expected loss of the k-nearest neighbor rule, conditioned on the data, converges to $(1 + 1/k)R^*$ for squared-error loss functions. For k_n-nearest neighbor rules where $k_n \to \infty$ and $k_n/n \to 0$, conditions are given under which the rules are asymptotically optimal.

In addition, methods of estimating the conditional risk of a rule with a particular data set are investigated. For a class of rules called local rules, the performance of two different estimates of the risk is bounded independently of the underlying distribution of (X,θ). This enables the statistician to construct confidence intervals for the risk of the rule and data he is using, without knowledge of the distribution of (X,θ).

ROBUST RANDOM PARAMETER ESTIMATION AND MINIMUM FISHER INFORMATION, R. Doraiswami (Coppe - Universidade Federal Do Rio De Janeiro, Brazil). An estimation of random parameters from a linear measurements model is proposed, when the a priori statistics are incomplete and only a small number of data points are available. The probability distributions of the random variables are assumed unknown. The only available statistics are the covariance of the random parameters and the 'partial covariance' of the measurement random variables. Using min-max theory, the optimal estimator is shown to be a soft-limiter.

ROBUST WIENER FILTERS, Tong Leong Lim and Saleem A. Kassam (Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104). Linear filtering of signals in additive noise is considered when the power spectral densities (psd's) of the signal and noise processes are not completely specified. Two models for classes of the psd's are defined, the ϵ-model for contaminated nominal psd's and the band-model with upper and lower bounds on the psd's. For the minimum mean-squared-error criterion, filters which are saddlepoint solutions for performance over the two classes of psd's are obtained; the saddlepoint or most robust filters are optimum for least favorable pairs of signal and noise psd's which have shapes tending to make them look alike. An example is presented illustrating the nature of the robust solution and the least favorable psd's, and showing the usefulness of the most robust filter in maintaining its performance over classes of psd's.

ON COUNTABLY INFINITE HYPOTHESIS TESTING, Jack Koplowitz (Department of Electrical and Computer Engineering, Clarkson College of Technology, Potsdam, NY 13676). Let $x_1,x_2,\ldots$ be a sequence of independent identically distributed Bernouilli random variables with unknown parameter $p = Pr\{x_i=1\}$. Let $S = \{p_1,p_2,\cdots\}$ be a countable subset ε $[0,1]$. Consider the countable set of hypotheses $H_i : p=p_i$, $i=1,2,\ldots$ together with the null hypothesis $H_0 : p \notin S$. Cover has demonstrated a decision procedure which makes only a finite number of mistakes with probability one in determining the true hypothesis for any $p \varepsilon [0,1] - N_0$, where N_0 is a set in $[0,1] - S$ of Lebesgue measure zero. The procedure generalizes to determining the mean with unknown distribution but finite variance.

Consider the hypothesis test H_0 : p is irrational vs. H_1 : p is rational. It has previously been shown that for any procedure which converges in probability under H_1, there exists an uncountable set of irrationals N_0, such that if $p \varepsilon N_0$ a zero limiting probability of error cannot be achieved.

On the other hand consider $S = \{p_1,p_2,\ldots\}$, $p_i = 1/2^i$. It can be shown that the hypothesis test H_i : $i=1,2,\ldots$, vs. H_0 : $p \notin S$ can be resolved for all $p \varepsilon [0,1]$. In this paper we obtain necessary and sufficient conditions for which the countable infinite hypothesis test can be resolved for all $p \varepsilon [0,1]$. It is shown that the true hypothesis can be resolved for all $p \varepsilon [0,1]$ if and only if the closure of the set S is countable.

A TEST FOR STATIONARITY, David J. Thomson (Bell Laboratories, Whippany, NJ 07981). A test for stationarity of a time series is given which has the properties of excellent performance in practice, simplicity, and modest computational requirements. The test procedure consists of three steps: Direct estimation of the spectrum on subsets of the data, comparison of the subset spectrum estimates at each frequency using the Bartlett M statistic and finally, testing the values of M for conformance to their null distribution using the one-sided Kolmogorov stastic D^-. This procedure is sensitive to changes localized either in time or frequency, and also indicates when excessively short subsets are chosen. It is shown that modest gains in sensitivity may be achieved by prefiltering the data but that smoothing the spectral estimates increases the probability of type I errors. A useful approximation for the power of the test is described and examples are given for both stationary and nonstationary series.

FISHER INFORMATION OF ORDER S, D.E. Boekee and Y. Boxma (Laboratory for Information Theory, Department of Electrical Engineering, Delft University of Technology, 4 Mekelweg, Delft, The Netherlands). The Fisher information is a well-known measure for the information about a parameter which is contained in observations.

It satisfies some basic properties which any information measure should satisfy. Its main application lies in estimation theory since the Fisher information can be used to obtain a lower bound on the accuracy of parameter estimators in terms of their variance. Furthermore, it is known that minimum variance estimator for a location parameter must have a gaussian distribution, in which case the Fisher information is minimized. However, it is well-known that the variance is not always a suitable criterion for the performance of estimators.

In this paper we consider a generalization of the Fisher information, called the Fisher information of order s. It is based on arbitrary powers of the absolute value of the derivative of the log likelihood function. This includes the Fisher information, if we consider second powers.

We shall consider some basic properties of this measure from an information theoretic point of view. It will be related to some other information measures.

It will be shown that this Fisher information can be used to obtain bounds on the s-th absolute central moment of parameter estimators, with $s \geq 1$. We introduce the notion of MSB estimators and obtain a class of density functions, called the exponential power distribution, for which MSB estimators contain minimal information. The bound obtained will be compared with other bounds on the s-th absolute central moment which are known in the literature. We shall also give results for the case that the parameter to be estimated is a random parameter.

We shall consider location and scale parameters as special cases and obtain some numerical results for the Fisher information of order s. It is possible to extend the notion of entropy power to the entopy moment of order s. We shall show that there exists an interesting analogy between the Fisher information of order s and the entropy power of order s.

Finally we shall discuss some applications of the Fisher information or order s to parameter estimation problems in additive noise.

SESSION C2

DATA SMOOTHING VIA ORDER STATISTICS, Shu-gwei Tyan (Polytechnic Institute of New York, Brooklyn, NY 11201). Through local monotonic regression an example is used to illustrate the optimality of some simple median-based smoothers. Properties of running medians, the building blocks of median-based smoothers, are explored by studying their fixed points. They are found to belong to two classes: the first class contains the locally monotonic sequences and the second contains only duo-valued sequences. Combinations of running medians of various lengths, in parallel and in serial, are also studied. Locally monotonic sequences which come naturally with running medians are shown to be closely related to other smoothers based on order statistics. If the upper and lower envelopes of a data sequence are defined as the local minmax and the local maxmin respectively, then it can be shown that the lower envelope of the upper envelope is locally monotonic. The same is true for the upper envelope of the lower envelope. Furthermore, the running median of the proper length always lies in between the two locally monotonic envelopes, thus they can be taken intuitively as the upper and the lower estimates.

SESSION C3

FREE DISTANCE PROPERTIES FOR A NEW CLASS OF TRELLIS PHASE CODES, J.B. Anderson and D.P. Taylor (Department of Electrical Engineering and Communications Research Laboratory, McMaster University, Hamilton, Ontario, Canada, L8S 4L7). Consider a constant-amplitude phase-varying signal of the form $\sqrt{2E/T}\cos((\omega_c+\Omega_n)t+\theta_n)$, $(n-1)T<t\leq nT$. Ω_n is a frequency offset chosen subject to $0\leq\Omega_n T<2\pi<<\omega_c T$, constant during any T-length interval. θ_n is chosen to make the (n+1)-st frequency change continuous in phase. By varying the offset frequency in response to an information sequence, one instruments a code in phase and time, which can provide a much reduced error rate over ordinary PSK and FSK in the same (or less) bandwidth. We present a useful class of such codes, the "multi-h" codes, and analyze their error performance. In these codes, the carrier frequency offset is either $2\pi\ell/q$ or 0, depending on the information bit, where ℓ and q are integers. The phase transitions possible with such a code form an endlessly repeating trellis phase structure; as such, only a limited number of phases and phase transitions must be recognized and demodulated during decoding, and a simple Viterbi algorithm scheme may be used. We calculate the free distance of all useful codes, that is, codes with constraint lengths 1 - 4, which utilize 20 or fewer different phases. Two-dimensional dynamic programming must be used, because the distances to phase trellis neighbors depend strongly on the information sequence. To verify these results, and to study burst errors, we have tested a software-implemented encoder/decoder on a simulated Gaussian channel. A 3-4 dB coding gain was found. Comments on spectral occupancy and a comparison to soft-decision decoding conclude the paper.

AN ANALYSIS OF SEQUENTIAL DECODING BASED ON CODE DISTANCE PROPERTIES, P.R. Chevillat, (IBM Research Laboratory, CH-8803 Rüschlikon, Switzerland), and D.J. Costello, Jr., (Department of Electrical Engineering, Illinois Institute of Technology, Chicago, IL 60616). The computational effort and the error probability of sequential decoding with fixed (time-invariant) convolutional codes are analyzed without employing any of the traditional random coding arguments. An upper bound on the computational distribution $P(C_t > N_t)$ for a specific fixed code is presented which decreases exponentially with the code's column distance. It is proved that rapid column distance growth minimizes the decoding effort and therefore also the probability of decoding failure (erasure). In an analogous way the undetected error probability of sequential decoding with a specific fixed code is proved to decrease exponentially with the free distance and to increase linearly with the number of minimum free-weight codewords. The two theorems prove that code construction for sequential decoding should maximize column distance growth and free distance in order to guarantee fast decoding, a minimum erasure probability, and a low undetected error probability.

SESSION C3

BRANCHING PROCESSES AND SEQUENTIAL DECODING, David Haccoun (Department of Electrical Engineering, Ecole Polytechnique de Montreal, Montreal, Canada). Modelling the incorrect paths explored by a sequential decoder as a special branching process with two absorbing barriers, this paper presents a procedure to bound the average number of computations per decoded bit. Using the same model a bound on the average population of the live incorrect paths existing at any depth away from the correct parent node will also be presented. These bounds are expressed in terms of the set of incorrect branch metrics actually used by the decoder, and in terms of the set of the stationary probabilities of Massey's Markov chain model of the metric differences on the correct path.

MODIFIED VITERBI DECODER FOR BURST CHANNELS, Jose Roberto B. de Marca and R.A. Scholtz (Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90007). This paper considers the decoding of convolutional codes transmitted over a finite-state burst channel, by applying Viterbi's algorithm to path estimation in a combined encoder-channel trellis diagram. Computer simulation is used to compare performance of this modified algorithm to a standard Viterbi algorithm, for several choices of two-site burst channels and convolutional codes. An upper bound to the performance of the modified decoder can be developed with the aid of flow-graph techniques, while simulation results for a "magic genie" decoder supply lower bound information for all possible decoders operating in this situation.

Complexity of the modified decoder grows linearly with the number of states in the channel model. However, the magic genie decoder does not have this problem since information concerning the true channel state is supplied to it by an external source (the "genie"). Replacement of the genie by a channel-state estimator is possible. This provides an interesting way to link interleaved decoders, for the purpose of sharing burst location information.

HADAMARD TRANSFORM-HAMMING DISTANCE DECODING RULE FOR CONVOLUTIONAL CODES, Ledinh Chon Tam (Centre de Recherche Industrielle du Québec, 245 Boul. Hymus, Pointe-Claire, Québec, Canada) and Jean-Pierre Adoul and Roger-Y. Goulet (Communication Research Center, University of Sherbrooke, Sherbrooke, Québec, Canada). A new recursion procedure is derived for the decoding of binary concolutional codes which minimizes the state Hamming distance. The procedure hinges on the use of Hadamard characteristic function in the Bayes estimation context. The structure of the decoder is shown to be fully modular and expandable and is illustrated for rate 1/2 and 2/3 "good" codes. Simulation performed on the binary symmetric channel shows that for fixed delay, the performances of this decoding rule is slightly

better than that of the Viterbi algorithm. For practical situations the decoder can be implemented without "true" multiplications and it is therefore suited for short codes with middle rates (e.g.: 1/2, 2/3, 3/4) specially when hard decision is necessary as in high speed applications.

PUNCTURED R = (n-1)/n CONVOLUTIONAL CODES FOR SIMPLIFICATION OF MAXIMUM LIKELIHOOD DECODING, J. Bibb Cain, George C. Clark, Jr., and John M. Geist (Harris Electronic Systems Division, Melbourne, FL). When rate -1/n codes are decoded using the Viterbi algorithm, each code state has 2 paths entering it for comparison of the respective metrics. However, with rate -(n-1)/n codes (n≥3) there are 2^{n-1} paths entering each state. Therefore, the resulting comparison and selection of the path with the best metric is made much more difficult. This will require either much more complex hardware in a high speed, fully parallel machine or more comparisons in a low-speed, serial machine. However, this problem can be avoided entirely by a judicious choice of code generator polynomials. By doing this one can decode the code just as one would decode a rate 1/2 code with very little additional complexity. Moreover, this technique provides a straightforward way of building a selectable rate decoder. The codes are constructed by periodically puncturing suitable lower rate codes, e.g., a R = 2/3 code can be obtained by deleting every fourth bit at the output of a R = 1/2 encoder. The resulting R = 2/3 code could be decoded as the original R = 1/2 code with unreliable decisions (or erasures) inserted in place of the deleted bits. The implementation advantages of the punctured code approach are discussed. In addition, we show Plotkin-type upper bounds on free distance for this class of codes and compare them with similar bounds for general R = (n-1)/n codes. A search for the best R = 2/3 (to k=10) and 3/4 (to k=11) codes of this type has been made, and the best codes are tabulated. It is shown that at these rates and constraint lengths one can almost always find a punctured code with a minimum free distance as good as the best R = 2/3 and 3/4 codes. Performance curves for these codes are also given.

SOME TRANSPARENT CONVOLUTIONAL CODES WITH A SIMPLE ENCODER INVERSE, Erik Paaske (Institute of Circuit Theory and Telecommunication, Technical University of Denmark, DK-2800 Lyngby, Denmark) and Rolf Johannesson (Department of Automata and General Systems Sciences, University of Lund, S-220 07 Lund 7, Sweden). As a counterpart to quick-look-in (QLI) convolutional codes, which are not transparent, we introduce rate r = 1/2 easy-look-in-transparent (ELIT) codes with a feedforward inverse (1+D,D). We present an extensive list of ELIT codes with an optimum generalized distance profile $\underline{d}_{\nu-2}$ =

$[d_0, d_1, \ldots, d_{\nu-2}]$, where ν is the constraint length and d_j is the order j column distance. In general ELIT codes have d_∞ superior to that of QLI codes.

SESSION C4

DIGITAL CODING OF ANALOG WAVEFORMS, Herbert Gish (Bolt Beranek and Newman, Inc., 50 Moulton Street, Cambridge, MA 02138). The digital coding of analog waveforms is considered from two aspects. One aspect is the evaluation of theoretical performance limits and the other is the design of efficient coding systems. Primary consideration is given to waveform transmission over the bandlimited, additive noise, Gaussian channel.

It is shown how rate distortion theory can be employed to obtain useful performance bounds on systems which are required to operate over a range of channel-to-noise ratios. The results obtained can provide information concerning the variation of system performance and, in particular, give quantitative information about the threshold effect. In addition a method of waveform coding is investigated which takes into account the differences in significance of the bits used in the binary description of the quantized samples of the waveforms. The performance of this technique is compared to theoretical limits and to PCM performance. Consideration is given to the tradeoff between threshold extension and the performance at high signal-to-noise ratios.

ASYMPTOTICALLY OPTIMAL BLOCK QUANTIZATION, Allen Gersho (Bell Laboratories, Murray Hill, NJ 07974). In 1948 W.R. Bennett used a companding model for nonuniform quantization and proposed the formula

$$D = \frac{1}{12N^2} \int p(x)[E'(x)]^{-2}dx$$

for the mean-square quantizing error when N is the number of levels, $p(x)$ is the probability density of the input, and $E'(x)$ is the slope of the compressor curve. The formula is an approximation based on the assumption that the number of levels is large and is a useful tool for analytical studies of quantization.

This note gives a heuristic argument generalizing Bennett's formula to block quantization where a vector of random variables is quantized. The approach is again based on the asymptotic situation where N the number of quantized output vectors is very large. Using the resulting formula, an optimization is performed leading to an expression for the minimum quantizing noise attainable for any block quantizer of a given block size k. The result specializes to known results for the one-dimensional case ($k=1$) and for the case of infinite block length ($k\to\infty$). The same heuristic approach also gives an alternate derivation of a bound of Elias for multidimensional quantization.

DIFFERENTIAL ENCODING FOR BINARY MARKOV SOURCES, George Thomas (Indian Institute of Science, Bangalore, India). A binary first order autoregressive process and its underlying memoryless state-transition process have the same rate-distortion function with respect to the probability-of-error criterion, for low values of distortion. In this paper we explore the possibility of encoding the memoryless process at a suitable rate R and distortion D and of deriving therefrom an (R,D)-approximation to the output of the Markov source. The motivation is that it is then possible to apply the techniques for source encoding of memoryless sources directly to the case of autoregressive sources driven by memoryless processes. It is immediately seen that the direct expansion of the encoded memoryless process using a feedback shift register circuit leads to unbounded error propagation. We propose a simple modification to the source encoder which generates error patterns which do not propagate. The rate-distortion performance attainable with the modified encoder is evaluated.

RUN-LENGTH ENCODING WITH FIDELITY CRITERION FOR BINARY MARKOV SOURCES, P.K.S. Wah (Swiss Federal Institute of Technology, ETH-Zentrum, 8092 Zurich, Switzerland). Several simple source encoding algorithms for binary stationary sources, both symmetric and asymmetric, with memory are described. One method is to sample one of every M (M>1) symbols and then to code these selected symbols with run-length code. At the receiving end the missing symbols are reconstructed according to the dependence between neighboring symbols. The resulting distortion for Markov sources of different orders can readily be calculated. Other methods include the exchanging of 1's to 0's or vice versa in the binary data to form long runs, so that isolated symbols, short runs or some certain run-lengths are eliminated. After this kind of artificial distortion process, the expected distortion d, the new run-length distributions of black and white runs $p(R_1=n)$ and $p(R_0=n)$ as well as the entropy H_r (minimum transmission rate) can also be given in closed form. The corresponding values of rates and distortions for different Markov sources are compared with the lower bound $R_L(d) = \frac{a}{a+b} H(a) + \frac{b}{a+b} H(b) - H(d)$, where H(x) represents "$- x\cdot\log(x) - (1-x)\cdot\log(1-x)$" and a and b are the conditional probabilities p(0/1) and p(1/0) of the binary source symbols. For many points they are quite close together. Computer simulations yield also the same results.

SESSION C4

DIFFERENTIAL PULSE-CODE MODULATION OF THE WIERNER PROCESS, Akira Hayashi (Kanazawa Institute of Technology, Kanazawa, Japan). The performance of differential pulse-code modulation with the Wiener process input is analyzed. The analysis is exact: no use is made of the concepts of slope overload error and granular error. The limit as $n \to \infty$ for the characteristic function of error distribution at time nT is found in terms of the step size parameter Δ and the number of levels N of quantization. Uniform symmetric quantization and ideal integration in the feedback path are assumed. Curves are presented of two kinds of mean-squared error versus Δ or N, and are compared with the rate distortion function.

THE SIMULATION PROBLEM, Joseph Linde and Robert M. Gray (Information Systems Laboratories, Stanford University, Stanford CA 94305). Let X be a stationary-ergodic source, U an i.i.d source with finite alphabet A consisting of M equally probable output levels and let $\rho(.,.)$ be a single letter distortion measure. Consider the class of all time-invariant, possible nonlinear, filters that can be represented as a shift register consisting of N M-level cells followed by a fixed function $F:A^N \to R$.

The simulation problem is to find, within this class, the filter which when driven by U will yield an output which is closest in the generalized Ornstein or $\overline{\rho}$-distance to the given source X. The finite entropy output process is called a fake process.

Gray had shown that the simulation problem is equivalent to the source coding (data compression) problem; in particular a good fake process filter can be used as a decoder in a source coding system, where a trellis or tree search is used as the encoder, to yield nearly optimal systems.

We use this approach to design source coding systems for i.i.d., autoregressive and moving average sources using fake processes generated via Central Limit Theorem and inverse distribution-scrambling functions techniques. Computer simulations are carried out for the case of Gaussian sources at the low rate of 1 bit/symbol. In this case it is shown that the systems so designed outperform the optimal quantizer by 0.7 dB in the i.i.d case and outperform delta-modulation and predictive quantization by 1-2 dB in the first order autoregressive and moving-average cases.

SESSION C5

SOME RELATIONS BETWEEN MUTUAL INFORMATION, SAMPLE PATH PROPERTIES, AND SIGNAL DETECTION, Charles R. Baker (Department of Statistics, University of North Carolina, Chapel Hill, NC 27514). Let (S_t) and (N_t), t in $[0,T]$, be stochastic processes having almost all sample paths in $L_2[0,T]$. (N_t) is assumed to be zero-mean and Gaussian. Under appropriate assumptions on (S_t), various conditions are known for finite mutual information of signal and signal-plus-noise, non-singular signal detection, finite signal-to-noise ratio (quadratic-linear test statistic), and signal sample path behavior (in terms of the noise covariance function). We present results which show the connection between these properties, for a fixed signal process (S_t) and a fixed noise process (N_t). When signal and noise are jointly Gaussian, the results given here generalize previous results due to T.S. Pitcher and to the author.

BOUNDS ON THE MUTUAL INFORMATION FOR NONLINEAR OBSERVATION PROCESSES WITH ADDITIVE WHITE GAUSSIAN NOISE, S. Arimoto and T. Hashimoto (Faculty of Engineering Science, Osaka University, Toyonaka, Osaka 560, Japan). This paper is concerned with the mutual information between an input signal $x(\cdot)$ and an output observable $y(\cdot)$ for linear or nonlinear observation processes of type, $y(t)=g(x(t),t) + w(t)$, where $w(t)$ is an independent white Gaussian noise. Lower bounds on the mutual information obtained for discrete-time processes are expressed by

$$I[x(0,k);\ y(o,k)] = \frac{1}{2} \sum_{i=0}^{k} \ln[\det(W_i^{-1}(W_i+P_i))]$$

where W_i are covariance matrices of the white Gaussian noise and P_i are certain non-negative definite matrices. One of upper bounds is described by

$$I[x(0,k);\ y(0,k)] = \frac{1}{2} \sum_{i=0}^{k} \ln[\det(W_i^{-1}(W_i+Q_i))].$$

In general, $Q_i = P_i$. If $g(x,t)$ is linear in x and $x(\cdot)$ is Gaussian, then $Q_i=P_i$ for all i. The bounds on the mutual information for continuous-time cases are described in terms of Fredholm determinant associated with certain linear integral equations, which are derived by taking limits of discrete-time processes as sampling period tends to infinitesimally small, provided that the continuous-time process $x(.)$ is mean-square continuous or continuous in probability.

The method proposed is effective for not only causal filtering but also non-causal filtering. Some new formulae for the mutual information $I[x(\tau,\theta);\ y(s,t)]$, where $s \leq \tau \leq \theta \leq t$, are presented.

STOCHASTIC AND MULTIPLE WIENER INTEGRALS FOR GAUSSIAN PROCESSES, Steel T. Huang (Department of Mathematics, University of Cincinnati) and Stamatis Cambanis (Department of Statistics, University of North Carolina at Chapel Hill). Multiple Wiener integrals and stochastic integrals are defined for Gaussian processes, extending the related notions for the Wiener process. It is shown that every L_2-functional of a Gaussian process admits an adapted stochastic integral representation and an orthogonal series expansion in terms of multiple Wiener integrals. Also some results of Wiener's theory of nonlinear noise are generalized to noises other than white, and the differential formula for the stochastic integral is derived.

A TIME PERTURBATION OF GAUSSIAN STOCHASTIC PROCESSES AND SOME APPLICATIONS TO THE THEORY OF SIGNAL DETECTION, A.F. Gualtierotti (Departement de mathematiques, Ecole Polytechnique Federale, 26, Av. de Cour, 1007 Lausanne, Switzerland). Let X be a stochastic process with square integrable paths. The statistician can in principle observe H(X), the closure in L_2 of the vector space generated by X. He is however interested in $L_2(X)$, that is, the family of "functionals" of the process (i.e. the family of functions measurable with respect to the tribe generated by X and square integrable). When X is Gaussian, one has

$$L_2(X) = \sum_{n\geq 0} H(X)^{\Theta n},$$

where "Θn" denotes the symmetric tensor product of n copies of H(X). Theoretically then the statistician has access to $L_2(X)$. In practice however information about X is obtained through experimental observations, which in turn lead to estimated parameters. Decisions are then based on the latter. It is thus important to evaluate how the above relation between H(X) and $L_2(X)$ is affected when one deals not really with X, but some perturbation of it, say Y.

The case of Y spherically invariant has been extensively investigated. In the talk we shall consider a time "perturbation" of X, i.e. set $Y_t = X_{At}$, and study some of its effects on detection and estimation. The two conclusions that may be derived are that nonsingularity of detection is not unduly affected, but that the usual procedures may introduce a systematic bias in estimations. The results are indicative of what may happen, rather than what really happens, because calculations about Y are made supposing X known.

ON THE RECONSTRUCTION OF THE COVARIANCE OF STATIONARY GAUSSIAN PROCESSES OBSERVED THROUGH ZERO MEMORY NONLINEARITIES, Stamatis Cambanis (Department of Statistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514) and Elias Masry (Department of Applied Physics and Information Science, University of California at San Diego, La Jolla, CA 92093). We consider the problem of

reconstructing the normalized covariance function R(t) of a zero mean stationary Gaussian process observed through a zero memory nonlinearity f(x), when we know the nonlinearity and the correlation function or the second order distribution of the output process. Three kinds of results are established showing (i) how arbitrary covariances may be reconstructed for certain nonlinearities: included here are monotonic f's, appropriate interval windows and certain quite general f's; (ii) how certain covariances can be reconstructed for arbitrary nonlinearities: included here are positive covariances (≥ 0), covariances with rational spectral densities, and bandlimited covariances; and (iii) how certain covariances can be reconstructed for certain nonlinearities: included here are covariances satisfying certain rather weak conditions, which are easily checked in terms of the output correlation function, and certain symmetric as well as nonsymmetric f's.

BOUNDS ON ESTIMATION ERROR FOR GAUSS-POISSON PROCESSES, Adrian Segall (Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa, Israel). Upper and lower bounds for the estimation error of a linear Gauss-Markov process that modulates the mean of a Gauss-Poisson process are obtained. The upper bounds are derived by looking at the best linear estimator for the same problem. The lower bound is obtained by using a generalized Cramer-Rao bound developed by Bobrovski and Zakai.

POISSON SAMPLING AND SPECTRAL ESTIMATION OF CONTINUOUS-TIME PROCESSES, Elias Masry (Department of Applied Physics and Information Science, University of California at San Diego, La Jolla, CA 92093). We consider a class of spectral estimates of a continuous-time stationary stochastic process $X = \{X(t), -\infty < t < \infty\}$ based on a finite number of observations $\{X(t_n)\}_{n=1}^{N}$ taken at Poisson sampling instants $\{t_n\}$. We investigate the bias and covariance structure of the estimates and discuss the influence of the spectral windows and the sampling rate on the performance of the estimates. The estimates are shown to be mean-square consistent under mild smoothness condition on the spectral density. The process X need not be bandlimited. The results of a simulation study are discussed.

PLENARY SESSION C

SOME RECENT RESULTS ON CHARACTERIZATION OF MEASURES OF INFORMATION RELATED TO CODING, J.D. Aczel (Mathematics Department, University of Waterloo, Waterloo, Canada). Essential properties of entropies and other measures of information are arrived at on basis of the noiseless coding theorems, source entropies and Huffman codes.

Conversely, characterization theorems are given, based on these properties. Some of the most important are Shannon's inequality, boundedness on an interval, subadditivity, additivity, branching and expansibility. Also entropies of mixed probabilistic and non-probabilistic character and convex f-divergences are dealt with, among others.

Some unsolved problems are stated.

TEAM DECISION, MARKET SIGNALING, AND INFORMATION THEORY, Y.C Ho (Harvard University, Cambridge, MA). This talk will attempt to unify three at least superficially separate fields of research: team decision theory, market signaling in economics and Shannon's information theory.

INFORMATION THEORY IN PHYSICS, E. T. Jaynes, Washington University, St. Louis, Mo. (USA). This talk will survey the role of information theory in physics, past, present, and future. There are a few interesting confluences of mathematical results known to workers in statistical mechanics; for example, application of information theory in predicting the course of time-dependent irreversible processes led to a formalism mathematically isomorphic with Norbert Wiener's prediction theory. The same integral equations for the optimal predictor appeared, the kernel covariance functions now appearing as thermal expectation values of products of quantum-mechanical operators.

SESSION D1

DECODING RANDOM CODES WITH AN OPTIMUM THRESHOLD, B. G. Dorsch, (Institut für Nachrichtentechnik, DFVLR, D-8031 Oberpfaffenhofen, Federal Republic of Germany (BRD)). For random block codes a suboptimum decoding rule is defined, which takes any codeword within a fixed distance t from the received word as decoding decision rather than the most probable one. It will be shown by random coding arguments for the binary symmetric channel that with an optimum threshold t channel capacity can be approached by this decoding rule. Furthermore for rates close to capacity the error exponent E(R) of this decoding principle is as good as that of maximum likelihood decoding. Comparisons are made also with lower and upper bounds of the error exponent of bounded minimum distance decoding BMD showing E(R) is superior for high rates and/or very noisy channels. Applying the results to the additive white Gaussian noise channel with hard binary decisions, e.g. for codes of length $N = 1000$ and rate $\frac{1}{2}$, yields an error probability $\leq 10^{-5}$ with a signal to noise ratio $E_b/N_0 \leq 3.6$ dB. This is approximately 2 dB lower than that of BMD when the minimum distance approaches the Varshamov-Gilbert bound.

DIGITAL WHITENING TECHNIQUES FOR IMPROVING SPREAD SPECTRUM COMMUNICATIONS PERFORMANCE IN THE PRESENCE OF NARROW-BAND JAMMING AND INTERFERENCE, Frank M. Hsu and Arthur Giordano (GTE SYLVANIA, Electronic Systems Group, Eastern Division, 77 "A" Street, Needham Heights, MA 02194). Narrowband jamming and interference due to other users in spread spectrum communications systems can be effectively suppressed by using digital whitening. The received spread spectrum signal is assumed to consist of the sum of the SS transmitted signals, narrowband interference and thermal receiver noise. Both the transmitted signal and the thermal receiver noise can be assumed to have white spectral characteristics. The narrowband interference, on the other hand, is nonwhite over the SS signal band and in the short term is therefore coherent. This coherent component can be predicted by use of digital whitening techniques which can be implemented as a transversal filter. Since the noncoherent portion of the received signal resulting from white signal components is not predictable, the estimate formed can actually be interpreted as an estimate of the interfering signal component. The narrowband interference is then suppressed by subtracting the estimate from the received signal. As a result, an impressive improvement in receiver performance can be obtained. In this paper, the Weiner and maximum entropy filters are used for digital whitening. The merits of these two filters in different jamming and signaling situations are compared and are found to exhibit similar performance over a wide range of input signal-to-noise ratios.

RECOVERY OF SPREAD SPECTRUM CARRIER FUNCTIONS, Karl Heinz Annecke, (Institut fur Elektrische Nachrichtentechnik der Rheinisch-Westfalischen Hochschule Aachen, D5100 Aachen, Alte Maastrichter Str. 23, W. Germany). A system is described which generates a synchronous spread spectrum reference carrier for a coherent receiver from the received signal by a decision-directed process. In a coherent receiver it is necessary in synchronizing a reference carrier to the carrier of the received signal. The synchronization is mostly done by a phase locked loop. In systems which use a wide-band carrier the difficulties to synchronize the reference carrier grow with the bandwidth of the used carrier. In this paper the carrier is a product of subcarriers $c_{ip}(t)$; $c_{ip}(t)$ are bipolar periodically repeated functions with period T_i consisting of m_i subpulses of length t_o $\left[c_p(t) = \prod_{i=1}^{n} c_{ip}(t)\right]$. The transmitted information is a bipolar function consisting of "bits" of length T; T is the period of the carrier $c_p(t)$ $\left[T = t_o \prod_{i=1}^{n} m_i\right]$. In the receiver each subcarrier $c_{ip}(t)$ is computed from the received signal which is multiplied with an estimate of the information and an estimate of the product of all the other subcarriers $\left[\prod_{\substack{k=1 \\ k\neq i}}^{n} c_{kp}(t)\right]$.

The systems properties are described for the special case n = 1 and for the common case n > 1. The theoretically expected results are compared with computer simulations and hardware measurements.

ERROR PROBABILITY FOR AN INCOHERENT CHANNEL WITH PARTIAL BAND OR TIME NOISE AND A MISMATCHED DECODING STATISTIC, Steven Krich, (Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA 02173). The bit error probability for an incoherent additive Gaussian noise channel with noise power density permitted to change for each channel symbol transmission has been evaluated by Viterbi and Jacobs (1975). They assumed a receiver which computes the weighted sum-of-the-squared matched filter envelope detectors with weights determined by an accurate estimate of the channel noise. They showed that by use of diversity (signal repetition) or coding the error probability could be made to decrease exponentially with E_b/N_o where E_b is the energy per bit and N_o w/Hz is the average noise power density.

We consider here the degradation in performance which results from an inability to accurately measure the channel conditions. For instance, the receiver might set the weights based upon the measured noise in a time interval just prior to the expected signal. If the noise is quickly time varying, the weights will be wrong much of the time. In spite of this, with diversity or coding, the error

probability still decreases exponentially in E_b/N_o in the presence of partial time noise. At most an additional 3 dB in E_b/N_o is required to compensate for the incorrect weights. By comparison, the error probability decreases linearly in E_b/N_o if no attempt is made to measure the noise.

Also considered is a receiver which ignores very noisy filter outputs and sets the remaining weights based upon an accurate measurement of the channel noise.

ADJACENT CHANNEL INTERFERENCE IN A BINARY BANDPASS COMMUNICATION SYSTEM, Israel Korn (Department of Elec. Eng., Faculty of Eng. Science, Ben-Gurion University, Beer-Sheva, Israel) and Moshe Herzberg (Faculty of Elec. Eng., Technion-Israel Institute of Technology, Haifa, Israel). The effect of adjacent channel interference on the probability of error in a binary bandpass communication system with an integrating and dumping detector is investigated. Narrowband filters are assumed in the receiver of the main signal and transmitters of both main and interfering signals. Plots of probability of error as a function of signal-to-noise ratio in the main channel or as a function of carrier frequency difference between the main and interfering signals are presented, assuming that the filters are of the Butterworth type. These figures are helpful in selection of minimal frequency spacing of adjacent channels.

CHANNEL ESTIMATION AND DECODING IN A MULTIPATH ENVIRONMENT, Kenneth S. Schneider (Network Analysis Corp., Glen Cove, NY) and Terrence P. McGarty (Communications Satellite Corp., Washington, D.C.). In the area of mobile satellite communications, one of the serious limitations is the fading due to specular multipath. In an attempt to mitigate against its effect, large aperture tracking antennas have been used on ships. However, such antennas are impractical for most land vehicles and aircraft so an alternative is sought. Recent work by Schneider and McGarty has shown that a signal processing technique using the intersymbol interference model proposed by Forney and the Viterbi decoder can provide adequate communications performance in the presence of specular multipath.

A limitation of this technique was that it required that estimates of the specular multipath delay and reflections coefficient be obtained for the matched filter. This is often a difficult set of estimates to obtain so that a robust scheme was proposed. The robust scheme uses a known matched filter and develops an equivalent finite state machine model. For this case the tap gains were to be estimated. Using reliable tap gains, it was shown that the performance of this robust scheme was similar to that of the optimum scheme with perfect channel knowledge.

In this paper the authors develop a technique to obtain the desired channel tap gains. This is done by using a Kalman filter with a training sequence. Schemes of this type have been proposed by Lawrence and Kaufman, and by Godard. We have built upon the latter technique and extended it to the vector measurement channel used by the robust scheme. The paper develops performance results and discusses them in terms of optimal sets of training sequences.

STOCHASTIC CHANNELS AS GENERALIZED COMMUNICATION NETWORKS, David Middleton (127 East 91st Street, New York, NY 10028) and J. Raymond Breton (Naval Underwater Systems Center, New London, CT 06320). Conceptually it is perhaps well known that physical channels through which information-bearing signals are transmitted can, in a rather loose sense, be regarded as space-time filters or generalized networks, analogous to the temporal-only filters or networks of conventional circuit and control theory. What has not, however, been constructed quantitatively is the analogous theory of such distributed filters, or networks, taking into particular account their capabilities for signal degradation and noise generation, specifically in terms of the controlling physics of the channel and the random and deterministic inhomogeneities which are the critical factors influencing information transmission.

New results include various operational and perturbation-theoretical solutions for the generalized network response functions; causality conditions which are the space-time equivalent of the Paley-Weiner conditions of (linear) temporal filters; control theory formulations which provide general algorithms for simulation and computational evaluation; various preliminary results for space-time stability conditions; procedures for calculating any desired channel statistics in terms of an appropriate equivalent deterministic channel (or distributed network). The approach is applicable to both linear and nonlinear systems, and is interdisciplinary, involving such disciplines as control theory, information theory, stochastic processes, the physics of propagation and scattering, and computational methods and techniques. It is illustrated here in detail with explicit results pertinent to electromagnetic and acoustical transmission.

PROLATE SPHEROIDAL WAVE FUNCTIONS - THE DISCRETE CASE, David Slepian (Bell Laboratories and University of Hawaii). A discrete time series has associated with it an amplitude spectrum which is a periodic function of frequency. This paper investigates the extent to which a time series can be concentrated on a finite index set and also have its spectrum concentrated on a sub-interval of the fundamental period of the spectrum. Key to the analysis are certain sequences, called discrete prolate spheroidal sequences, and certain functions of frequency called discrete prolate spheroidal functions. Their mathematical properties are investigated in great detail and many applications to signal analysis are pointed out.

SESSION D2

INTRINSIC DIMENSIONALITY SPATIAL-TEMPORAL ARRAY PROCESSING, Salvatore D. Morgera (Raytheon Company, Submarine Signal Division,, Systems Laboratory, Portsmouth, RI 02871). This work is concerned with efficient spatial-temporal array processing for the purpose of detecting a target waveform in an unknown interference noise environment. A new and rather general characterization of the interference noise field spatial-temporal covariance structure is presented. The covariance model is felt to be applicable to the acoustic, radar, and seismic areas and conveys an appreciation of the fact that the intrinsic dimensionality, i.e., the number of linearly independent spatial-temporal parameters, of the interference noise field is generally much smaller than the physical or "full" dimensionality of commonly employed array processors. The structure and performance of a reduced dimensionality array processor is investigated, with the performance compared to both a full dimensionality spatial-temporal processor and to just a temporal processor. It is shown that the performance of the reduced dimensionality spatial-temporal processor can be highly competitive with that of the full dimensionality processor if (all but one of) the spatial "look" directions are chosen as the principal components of the interference noise vector stochastic process corresponding to the largest values of interference noise-target waveform spatial-temporal correlation. Since the principal components are not a priori known, a scheme is presented for estimating the "best" set of spatial directions. Apart from the obvious advantages of the reduced dimensionality processor in the way of minimum throughput and memory demands, the reduced dimension permits maximum convergence rate and, consequently, utmost sensitivity to a quasi-stationary interference noise environment for an adaptive implementation.

NONLINEAR FEATURE EXTRACTION WITH A CRITERION OF A GENERAL FORM, Keinosuke Fukunaga and Robert D. Short (School of Electrical Engineering, Purdue University, West Lafayette, IN 47907). In this paper the optimal nonlinear feature extraction for a criterion function of the general form $f(D_1,\ldots,D_M, K_1,\ldots,K_M)$ (where the D_j's and the K_j's are the conditional first and second order moments) is considered. The optimal solution is found to be a parametric function of the conditional densities. By imposing a further restriction on the functional dependence of f on the K_j's, the optimal mapping takes on an intuitively pleasing function of the posteriori probabilities. Next, the optimum feature mapping is restricted to a finite dimensional subspace. The resulting optimum mapping is a linear combination of the projections of the posteriori probabilities onto the subspace. The problem of finding the best single feature selection is discussed in this framework. Finally, several examples are also discussed.

SESSION D2

AN ALGORITHM FOR OPTIMAL NONLINEAR STRUCTURE PRESERVING FEATURE EXTRACTION, Scott A. Starks and Rui J.P. de Figueiredo (Rice University, Houston, TX 77001). This talk presents a new approach to nonlinear structure preserving feature extraction. This method is based on certain graph theoretical considerations (such as the minimal spanning tree, edge inconsistency, and diameter edges) and topographical considerations (such as interpoint distance measures). After introduction of the subject matter and appropriate background material, the algorithm is formulated.

Numerical results from the application of this algorithm to various test data sets are presented. Evaluation of these test results are quite encouraging.

RECURSIVE FACTOR ANALYSIS METHODS IN FEATURE EXTRACTION PROBLEMS, L. Kurz and C. S. Yoon (Polytechnic Institute of New York, 333 Jay Street, Brooklyn, NY 11201). The problems of data reduction, feature extraction and pattern recognition using factor analysis techniques were studied previously. The basic problem considered in this paper is to find the estimates of the factor loading matrix $\underline{A}$, and the unique matrix, $\underline{D}^2$, based on partitioned observations with the assumption that the underlying statistical parameters undergo slow change. This assumption corresponds to many useful data sets such as sleep-stage encephalograms, earth resources satellite data, etc. The main approach is to obtain the desired estimates from the partitioned data by maximum likelihood estimation of $\underline{D}^2$ in conjunction

with estimation of $\underline{A}$ by a stochastic approximation procedure. The approach used in this paper yields an efficient computer-oriented algorithmic approach to the feature extraction problem.

TEXTURE MODELING USING STOCHASTIC TREE LANGUAGES, S.Y. Lu and K.S. Fu (School of Electrical Engineering, Purdue University, West Lafayette, IN 47907). In the proposed stochastic model, a texture pattern is divided into fixed-size windows. A stochastic tree grammar (STG) is then used to characterize windowed patterns of the same class. The construction of an STG from a training texture pattern is as follows: A single pixel or a window of n x n pixels with a relatively uniform gray level is chosen to be the basic element of the texture pattern. Gray levels of basic elements are treated as pattern primitives. Then, a windowed pattern is transformed into a tree representation. Each basic element in the window corresponds to a node in the tree representation; hence, a pattern primitive is represented by a label assigned to its corresponding node. Consequently, we have a set of trees obtained from the training patterns. An STG can then be contructed from the set of trees to model the class of training patterns.

Texture synthesis, generation and discrimination using the proposed stochastic model are reported.

BEST FIRST PARSING OF NOISY WAVEFORMS, Laveen N. Kanal and George C. Stockman (Laboratory for Pattern Analysis, Department of Computer Science, University of Maryland, College Park, MD 20742). This paper presents a paradigm for linguistic analysis of noisy waveforms, which views analysis as simultaneous search of paths in the model space and the data space. A set of possible paths, commonly called segmentations, are explored by the primitive feature extractor. Each data path is of the form $\delta_i = d_{i,1}(P_{i,1})\ldots d_{i,n_i}(P_{i,n_i})$ where $d_{i,j} \in V_T$, the terminal vocabulary of the grammar and $P_{i,j} \in [0,1]$ is the corresponding confidence or probability of extraction. The purpose of the grammar model is to accept only those segmentations δ_i that have correct syntactic structure, and the goal of analysis is to develop all correct interpretations if necessary and in best-first order. A convenient evaluation for a correct interpretation is $f(\delta_i) = \min \{P_{i,j} : j = 1,n_i\}$.

All possible paths through a context free grammar $G = \langle S, V_N, V_T, P\rangle$ can be explored via state space search by starting with the initial symbol S and canonically generating all sentential forms. If $s = \beta\gamma$ is any sentential form generated, where $\beta = t,\ldots t_k$ is a string of terminals, then the degree of match between s and the data is defined as $\hat{f}(s) = \min \{P_j : j = 1,k\}$ where β is the initial segment of some segmentation $\delta_i = t_1(P_1)\ldots t_k(P_k)\, d_{i,k+1}(P_{i,k+1})\ldots d_{i,n_i}(P_{i,n_i})$. If β is not the initial segment of any segmentation, then

$\hat{f}(s) = 0$ and the search path is dead at state s. If $s = \beta = \delta_i$ for some segmentation, then β is a final state of the search and a correct interpretation of the data having confidence $f(\beta) = \min \{P_{i,j}: j=1,n_i\}$. Since $\hat{f}(s) \geq f(\beta)$ for any sentence β derivable from sentential form s, a best-first search of the space of sentential forms is an admissible algorithm for producing the highest confidence interpretation of the data. Because the search paradigm is so simple, it is easily extended to include bottom-up and non-left-right syntactic operators and extraction of primitives under syntax control.

SYNTACTIC SIGNAL PROCESSING, Francois Le Chevalier, Gérard Bobillot and Cécile Fugier-Garrel (Office National d'Etudes et de Recherches Aérospatiales (ONERA) 92320 Chatillon, France). A new representation of signals as a string of simple elementary signals whose succession obeys syntactic rules is presented; the extraction of such signals can then be performed in real time by using an original syntactic decoding method, DESCOA, effective for a wide variety of type 0 languages.

Extensions of this method lead to a true syntactic signal processing, including adaptive syntactic processing and syntactic filtering. Simulation results are shown for three examples: radar target identification, syntactic processing of frequency-modulated signals, and adaptive syntactic antenna processing.

VISUAL PERCEPTION, INVARIANTS, NEURAL NETS, H.D. Block, D.C. Lewis, and R.H. Rand (Department of Theoretical and Applied Mechanics, Cornell University, Ithaca, NY). One problem in the modelling of perception is to recognize that the same object is being observed, even when the visual representation of that object on the retina is subjected to various transformations. That is, we seek an output which is invariant under a class of transformations on the inputs.

We first formulate mathematically the classes of transformations to which the patterns will be subjected; e.g. translation, rotation, dilation, projection, irregular illumination, motion, distortion, etc. Next we consider various processing operators applied to the patterns; e.g. autocorrelation, Fourier transform, Laplacian, gradient, convolution and more general operators. The successive application of the pattern transformations and processing operations have an algebraic structure which we are studying, with particular interest in the invariants; i.e. those operators which are unchanged by a class of transformations applied to the visual patterns. Finally we present neural networks which implement the various processing operators, and we consider their biological plausibility.

SESSION D3

BOUNDS ON THE DELAY COMPLEXITY OF ERROR CORRECTING CODES, Yeshoshua Imber and John E. Savage (Program in Computer Science, Brown University, Providence RI 02912). The time to decode is important in high-speed data communications and in other applications including error-correction in computer memories. The delay complexity of decoders, which is a measure of decoding time, is examined and lower bounds are derived in terms of the rate and error-correction capability of the code used and the probability of error provided. Decoders and codes that achieve many of the lower bounds are described. In addition tradeoffs between code rate and decoding time are studied for three types of single error correcting codes.

TIME TO FIRST ERROR FOR AN INTERLEAVED CODE ON A BURST-NOISE CHANNEL, Robert A. Rutledge (IBM Corporation, D18 - 707-2, P.O. Box 390, Poughkeepsie, NY 12602). A bit-interleaved random error correcting code on a burst-noise channel is considered. The channel noise distribution is assumed to follow a simple Markovian model. The probability that a codeword is uncorrectable (PWE) was obtained in a previous paper. The present work gives tight bounds on the mean time to the first uncorrectable codeword (MTBF). The effect of the degree of interleaving on each of these parameters is illustrated with an example. It is shown that (for this example) a small degree of interleaving will increase MTBF, but also increase PWE.

A method for extending these results to the more interesting case of a block-interleaved, burst error correcting code is indicated.

LONG BLOCK CODES CAN OFFER GOOD PERFORMANCE, David Chase and H. David Goldfein (CNR Inc., 220 Reservoir Street, Needham, MA 02194). Binary algebraic decoders can be used directly with channel measurement (soft decision) information when algorithms within the class originally proposed by Chase are utilized. These algorithms are based on iterating a binary decoder by inverting received binary digits which are likely to be incorrect and selecting the most likely error pattern when more than one solution is found. In this paper a technique is introduced for using short block codes to reduce the number of iterations required to achieve a given level of performance when decoding long block codes with channel measurement information. The effectiveness of this approach is demonstrated by decoding an (N,K;D) = (128,64;22) BCH code with the use of a (23, 12;7) Golay code for perturbing the raw data which feeds the binary decoder. The simulated performance indicates that bit errors below 10^{-5} can be achieved at a signal-to-noise ratio per information bit, E_b/N_0, as low as 3.5 dB over an additive Gaussian noise channel. This is believed to represent the present state of the art for decoding rate-½ block or convolutional codes.

SESSION D3

IMPROVEMENTS IN BLOCK-RETRANSMISSION SCHEMES, John J. Metzner (Electrical and Computer Engineering Department, Wayne State University, Detroit, MI 48202). Consider the case where an (N,K) block encoded word cannot be decoded reliably and a second block of N redundant digits is sent to allow the receiver to make a new try based on the combined information received. (It is assumed that a likelihood ratio or at least a "soft" decision is retained for each digit.) Two classes of schemes are proposed and analyzed which give significantly better performance than is obtained by sending a repeat of the first block, yet do not require excessive decoding complexity.

One approach is to take small sub-blocks of the original N-digit code as the data digits of a short rate one-half code. For example, for a sub-block of 4 binary digits the second sending can complete an (8,4) Reed-Muller code with a minimum distance of 4. (With retransmission of the four digits, the minimum distance (original and repeat) would be 2). The likelihood ratio information can be used to make "hard" (binary) decisions for the four digits in the subcode (based on the (8,4) code structure.) To allow for the case where the first sending is garbled, the encoding should permit the data to be extractable from the second sending alone.

The other approach is to treat the first sending as the data digits of a systematic convolutional code of short constraint length. The likelihood ratios retained from the first sending can be used in Viterbi-decoding type procedures, with the original (N,K) code used for a final check. Constraint lengths as short as two or three digits can yield significant improvement over retransmission in some situations.

ANALYSIS OF MERGING IN THE BIT-BY-BIT DIRECTION/DECODING ALGORITHM, B.D. Fritchman and J.C. Mixsell (Department of Electrical Engineering, Building 19, Lehigh University, Bethlehem, PA 18015). Even though the Viterbi algorithm is not strictly sequential, part of its value lies in the fact that under most conditions, the computational process for the decision on the k^{th} symbol truncates (merges) at some step $k + D$, $D > L - 1$. The algorithm truncates quickly when the SNR is large, and more slowly when it is small.

We show that the sequential bit-by-bit algorithm also exhibits this property of merges, but its characteristics are the reverse of the Viterbi algorithm. If the algorithm has not merged at the constraint length, merging beyond the constraint length becomes slower as the SNR increases.

The quality of the decision is investigated as D increases by decreasing the decision for the k^{th} symbol and allowing the delay

constraint D to increase. It is shown that a value D* exits such that for values D > D*, no change in the decision error can occur.

A geometric signal space description of the conditions for a merge at any step is given. The conditions presented apply to the detection of signals in the presence of intersymbol interference, to the decoding of convolutionally encoded information, and to the joint detection/decoding problem.

IMPLEMENTATION OF DECODERS FOR BLOCK CODES, E.R. Berlekamp (University of California, Berkeley, CA). This talk will describe several decoders which the author recently designed.

A decoder for the (31,15) Reed-Solomon code had been implemented in standard T^2L technology, using a total of 78 16-pin integrated circuit packages. This decoder runs on a synchronous clock whose cycle time can be as short as 120 nanoseconds. The decoder can correct any pattern of t errors and s erasures for which $2t+s \leq 16$ in no more than 3559 clock cycles. This high speed is achieved by a novel architecture. No microprocessors are used.

Extrapolations of this performance to longer block lengths and larger alphabet sizes indicate that on quieter channels a much higher throughput can be achieved by using higher rate RS codes of much longer block lengths with only a small increase in hardware complexity.

We have also designed high-speed decoders for long <u>binary</u> codes using soft decisions.

Traditional analyses of the performance of Reed-Solomon decoders in the presence of both errors and erasures have assumed that decoding failure occurs only if there is no codeword within the appropriate distance of the received word. In applications where decoding error is much more costly than decoding failure, performance can be improved by introducing the notion of "false erasures". These are characters which are correct even though the demodulator erased them (presumably because one or more wrong signals, as well as correct signal, scored above some threshold). In a typical Gaussian noise environment, about 20 or 25 percent of the "erased" characters may actually be false erasures, and this fraction can be increased by modifying the demodulator's thresholds. Since this fraction of false erasures is significantly larger than q^{-1}, the false erasures can be used to provide an increased capability of detecting decoding errors. This capability is especially helpful on blocks which contain many erasures. For example, since the (31,15) Reed-Solomon code is maximum distance separable, any word containing 16 erasures can be decoded to a codeword, even though this codeword is correct only if the block contains no errors.

However, after the decoder has "corrected" the 16 erasures, a posterior check of its accuracy can be obtained by counting the purported number of "false erasures", e.g., the number of erasures in which the purported correction consisted of no change. If the block indeed contained no character errors, then the expected number of false erasures is 25 per cent, or 4 characters, but if the block contained one or more errors, and the decoding is therefore incorrect, then the expected number of false erasures is 1/32, or about 1/2 of a character.

When additional processing time is available, the decoder might also use the false erasures to correct many error patterns with $2t+s > 16$.

CODING FOR NONPROBABILISTIC, DISCRETE CHANNELS, William L. Root (The University of Michigan, Ann Arbor, MI 48109). Coding problems are considered for discrete channels with finite alphabets. The mappings from input to output are deterministic, but are a priori unknown to sender and receiver and may change from time to time. These mappings do, however, belong to a class of mappings known to the communicators. With error-free decoding the criterion of satisfactory transmission, maximum rates are estimated, or in a few instances calculated, for various classes of channel mappings. Most of the channels are memoryless; one is a special channel with finite memory.

CHAIN CODING OF TABULAR DATA IN NOISY ENVIRONMENTS, L. Kurz and C. Mohwinkel (Polytechnic Institute of New York, 333 Jay Street, Brooklyn, NY 11201). In this paper, some methods for reducing tabular data computed by noise into a string language by means of chain encoding are presented. The methods of detection of chain angle, chain construction, and the use of positional information of the input data for checks of position and accuracy of any individual chain element guarantee accurate description of the object by the chain codes. The data reduction methods presented in this paper are useful in increasing speed and efficiency of tabular (pictorial) data transmission and in reducing expensive storage and processing time of the received encoded image. In addition, chain coded data has certain indigenous properties which make it useful in certain types of recognition and correlation problems.

AD-A051 147 INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS INC--ETC F/G 9/4
INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY HELD OCTOBER 10-1--ETC(U)
1977 T BERGER, R E BLAHUT F49620-77-C-0128

UNCLASSIFIED AFOSR-TR-78-0195 NL

2 OF 2
AD
A051147

END
DATE
FILMED
4 -78
DDC

ROBUST RATE DISTORTION THEORY, Dimitri Kazakos (State University of New York, Buffalo) and T. Papantoni - Kazakos (Bell Laboratories and Rice University). The practical utility of Rate-distortion theory is hampered by the fact that exact knowledge of source statistics is necessary in order to seek the optimal source codes. Recently, several researchers developed source coding techniques for unknown source statistics. However, an unduly high complexity of implementation limits their practicality.

In the present paper we consider the adaptation of Rate-distortion theory to the situation of incompletely known source statistics. We assume that the source statistics are members of a convex family and we seek the worst source within the family. The explicit calculation for some examples is carried out.

NEW RESULTS ON CODING OF STATIONARY NONERGODIC SOURCES, Lee D. Davisson and Alberto Leon-Garcia (University of Maryland, College Park, MD) and David L. Neuhoff (University of Michigan, Ann Arbor, MI). We present two results on the coding of stationary nonergodic sources. The first is a converse information transmission theorem. The least achievable average distortion in transmitting the output of a stationary nonergodic source through a channel of capacity C using block stationary encoder/decoder pairs is given by $\delta(C)$. In general, $\delta(C) > D(C)$, the distortion-rate function. Gray and Davisson have shown that $\delta(R)$ is also equal to the least average distortion achievable by the fixed rate coding of stationary nonergodic sources. Our second result is a source coding theorem stating that there exist variable rate codes which achieve R(D), the rate-distortion bound, and hence also achieve D(R).

VARIABLE-RATE UNIVERSAL BLOCK SOURCE CODING SUBJECT TO A FIDELITY CONSTRAINT, K.M. Mackenthun, Jr., and M.B. Pursley (Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801). Previous results on universal variable-rate source coding subject to a fidelity constraint are given in terms of a probability measure W on a collection Λ of stationary ergodic sources. Loosely speaking, these results imply that there exist variable-rate codes which yield average distortion D at an optimum rate $R_\theta(D)$ for all sources θ in some subset $\Lambda' \subset \Lambda$. For arbitrary $\varepsilon > 0$, Λ' can be chosen to satisfy $W(\Lambda - \Lambda') < \varepsilon$.

New results are presented which do not require a probability measure on Λ. In particular, conditions are given under which there exist variable-rate codes whose rate-distortion performance converges to $R_\theta(D)$ for <u>each</u> source θ in the collection Λ. Two such forms of

universal coding are considered. For the weaker of these two (weakly-universal) the convergence is pointwise, and for the stronger (strongly-universal) the convergence is uniform on Λ.

Existence of weakly-universal variable-rate codes is established for a collection of stationary ergodic sources under very general conditions. One set of conditions is as follows: (1) the source and reproduction alphabets are separable metric spaces, (2) the distortion measure is a nonnegative, nondecreasing, continuous function of the metric, (3) there exists a reproduction letter giving finite average distortion for all sources in the collection, and (4) $\sup\{R_\theta(D): \theta \in \Lambda\} < \infty$. Other sets of conditions are given which, for instance, replace (4) by conditions on the memory of sources in the class. Examples of collections of sources which are covered by these other sets of conditions but do not satisfy (4) include the collection of all stationary ergodic Gaussian sources with rational spectral density.

Existence of strongly-universal variable-rate codes is established for collections of finite alphabet memoryless sources and certain types of Markov sources. For example, it is shown that universal codes exist for the collection of all binary first-order Markov sources, but they do not exist for the class of all finite-order Markov sources.

COMPRESSION OF INDIVIDUAL SEQUENCES VIA VARIABLE-RATE CODING, Jacob Ziv (Bell Laboratories, Murray Hill, New Jersey 07974) and Abraham Lempel (Sperry Research Center, Sudbury, MA 01776). Compressibility of individual sequences by the class of generalized finite-state, information-lossless encoders is investigated. These encoders can operate in a variable-rate mode as well as a fixed-rate one, and they allow for any finite-state scheme of variable-length-to-variable-length coding. For every individual, infinite sequence x we define a quantity $\rho(x)$, called the compressibility of x, which is shown to be the asymptotically attainable lower bound on the compression-ratio which can be achieved for x by any finite-state encoder. This is demonstrated by means of a constructive coding theorem and its converse which, apart from their asymptotic significance, also provide useful performance criteria for finite, and practical, data-compression tasks. The proposed concept of compressibility is also shown to play a role analogous to that of entropy in classical information theory, where one deals with probabilistic ensembles of sequences rather than with individual sequences.

CODING THEOREMS FOR INDIVIDUAL SEQUENCES, J. Ziv (Bell Telephone Laboratories, Inc., Murray Hill, NJ 07974). A quantity called the finite-state complexity is assigned to every infinite sequence of

elements drawn from a finite set. This quantity characterizes the largest compression ratio that can be achieved in accurate transmission of the sequence by any finite-state, fixed rate encoder (and decoder). Coding theorems and their converses are derived for an individual sequence without any probabilistic characterization, and universal data compression algorithms which are asymptotically optimal for all sequences over a given alphabet are introduced.

The finite-state complexity of a sequence plays a role similar to that of entropy in classical information theory (which deals with probabilistic ensembles of sequences rather than an individual sequence). For a probabilistic source, the expectation of the finite state complexity of its sequences is equal to the source's entropy. The finite state complexity is of particular interest when the source statistics are unspecified.

More recent results which lead to a distortion-rate theory for individual sequences will also be discussed.

ERGODIC AND MIXING PROPERTIES OF A CLASS OF COMPOSITE SOURCES, Robert J. Fontana (Department of Electrical Engineering, Stanford University, Stanford, CA). Sources whose probability distributions vary from letter to letter are known as composite sources. Such a source can be modelled as a combination of a collection of random processes and an additional "switch" random process which chooses an element from the collection at each instant of time. The ergodicity (weak-mixing, and strong-mixing) of the composite source is examined, and in doing so, a one-to-one correspondence is noted between a regular composite source and an infinte memory channel. Further, for regular composite sources, satisfying certain asymptotic independence constraints, the mixing properties of the source are determined by the mixing properties of the switch. Finally to better illuminate these principles, examples are given of composite sources satisfying the regularity and asymptotic independence properties.

SESSION D5

AN IN-PLACE SELF-REORDERING FFT, James K. Beard (Honeywell Inc., 5303 Shilshole Ave. NW, Seattle, WA 98107). The FFT algorithm inherently involves a requirement that the data be bit-reverse address reordered, either before or after the transform operation. If the algorithm is an in-place FFT, a separate pass is required to perform this operation. As a typical example, a 1024-point radix 2 algorithm requires five FFT passes and one reordering pass, and thus spends about 17 percent of its time in the reordering pass. The use of a separate reordering pass can be eliminated, so that no extra time is required for the reordering operation for an in-place FFT, by a technique using the properties of the reordering operation to perform an FFT pass.

SYNTHESIZING LARGE DFT'S USING THE MUTUAL PRIME FACTOR CYCLIC ALGORITHM, Salvatore D. Morgera (Raytheon Company, Submarine Signal Division, Systems Laboratory, Portsmouth, RI 02871). The utilization of a Mutual Prime Factor Cyclic Algorithm (MPFCA) for efficiently performing large Discrete Fourier Transforms (DFT) is examined. The mutual prime factor portion of the algorithm is originally attributed to L.H. Thomas, with generalization supplied by I.J. Good; the cyclic portion of the algorithm has been recently formalized by S. Winograd. Two methods are presented for implementing large size transforms in a range of interest from 60-5000 points. Computational complexity of the large transforms is estimated for both methods and compared with that of the FFT. The computational savings are found to be quite substantial - either approximately ($\frac{1}{2}$) or ($\frac{1}{4}$) the number of multiplications required by the FFT, depending on the method employed. These results bring to light the importance of a variable modulo arithmetic and addressing capability for high speed frequency domain processing for a multitude of important functions, e.g., modem design, spectral analysis, FIR digital filtering, beamforming, etc.

COMPUTATIONAL EFFICIENCY OF NUMBER THEORETIC TRANSFORM IMPLEMENTED FINITE IMPULSE RESPONSE FILTERS, Thomas A. Kriz and Dale F. Bachman (IBM Federal Systems Division, Owego, NY 13827). The computational merits of using Number Theoretic Transforms to spectrally implement finite impulse response (FIR) filters are discussed. It is shown that, for all but low tap filters, the spectral method yields a significant reduction of the multiply load and a modest reduction of the add load in comparison to the usual time domain methods.

SESSION D5

MULTIPLICATIVE CONVOLUTIONS AND FOURIER TRANSFORMS, S. Cohn-Sfetcu, (Bell-Northern Research, Ottawa, Canada). Mellin convolutions and Hankel transforms are analyzed as convolutions and Fourier transforms on multiplicative locally compact abelian groups, with a view towards providing fast algorithms for use in signal processing.

THE SOLUTION OF DISCRETE CONVOLUTIONS WITH A BOUNDED ERROR CONSTRAINT, A. Arcese (Electrical Engineering Department, Universidad de Los Andes, Bogota, Columbia). In this paper we present a fast new algorithm for solving discrete convolutions with a bounded error constraint. Unlike least squares, the algorithm does not require inverting or forming the elements of a matrix, but rather, operates directly on the data. The algorithm is demonstrated for the prediction equation on a synthetic speech waveform.

ERROR CONTROL IN ARRAY PROCESSORS, Dhiraj K. Pradhan (Department of Computer Science, University of Regina, Regina, Saskatchewan, Canada). Current LSI technology makes it feasible to have computational facility, built by using large array processors. It is not impractical to envisage thousands of microprocessors operating in an array.

This paper proposes a mode for error control in array processor, operating in single-instruction-multiple-data stream (SIMD) mode. In this mode of operation, each processing element computes a single function which is the identical function for all processors operating in the array.

The error-control technique proposed here is an extension of an earlier technique (for error-control in bit-wise logical operations). The technique proposed in this paper uses in a novel way generalized Reed-Muller Code (GRMC) for designing fault-tolerant array processor. This design encodes data streams into codewords in an i-th order GRMC. Then, these codewords are processed by an array processor which contains redundant processing elements.

It is shown in this paper that the resulting output vector forms a codeword in a higher order code. The output vector is then decoded for error-correction/detection and thus provides fault-tolerance.

DUAL-MODE LOGIC: A NEW REDUNDANCE TECHNIQUE FOR THE DESIGN OF EASILY TESTED LOGIC CIRCUITS, S. DasGupta (IBM Corporation, Fishkill, New York) and C.R.P. Hartmann and L.D. Rudolph (School of Computer and Information Science, Syracuse University, Syracuse, New York). Traditionally, logic circuits have been designed with little regard for the problem of fault testing. Designers

have been concerned with minimizing the complexity of the circuit and this was certainly understandable in the period prior to the development of LSI. When a logic circuit is small, there are a number of standard techniques for deriving tests to detect faults. As the size of the circuit grows, however, the number of tests required may increase very rapidly. Moreover, the complexity of finding the test set may become prohibitive. As logic component costs decrease and the difficulty of testing increases, a point is reached where logic designers should begin to consider the possibility of introducing hardware redundancy to simplify testing.

In this talk, we present one such approach based on the concept of dual-mode logic. The basic idea is to use redundancy (1) to reduce the number of tests required, and (2) to reduce the complexity of deriving the tests. From a testing standpoint, the ultimate solution would be a method of designing logic networks which could be tested using a function-independent test set of minimum possible size, i.e., a test set that is independent of the logic function and internal layout of the network. For stuck-at faults in combinational networks this is achieved through the use of a family of dual-mode gates which perform the required logic functions in one mode and are tested in the other mode. It is shown that in combinational networks built with dual-mode gates, all patterns of t or fewer stuck-at faults can be detected with only two function-independent tests provided that the combinational network has $d = 2t+2$ extra inputs. In sequential networks built with dual-mode gates, all patterns of t or fewer stuck-at faults can be detected with only six function-independent tests provided that the sequential network has $d = 2t+4$ extra inputs.

GROUP THEORETIC CODES FOR BINARY ASYMMETRIC CHANNELS, Serban D. Constantin and T. R. N. Rao (Department of Computer Science, Southern Methodist University, Dallas, Texas 75275). With the advent of the large scale integrated memory technologies, unidirectional failure properties have been observed in the functional behavior of the new memory cell: a cell failure is more likely to be a '1 → 0' transition (or 1-error) than a '0 → 1' transition (or 0-error) whose probability of occurrence is low. In the ideal case, the probability of an 0-error can be assumed 0.

Starting with the asymmetric distance $d_a(X,Y)$ between two binary n-tuples X and Y, a metric previously defined by one of the authors, and using the minimum asymmetric distance requirement for a single asymmetric-error (in our case 1-error) correcting code, we introduce a new class of codes suitable for the ideal binary asymmetric channel which we will refer to as 'group theoretic codes'. The single asymmetric-error correcting capability of these codes is proved and their superiority in information rate over previously developed single error correcting codes for the ideal binary asymmetric channel or the binary symmetric channel is established.

Given an Abelian group G(+) with elements $a_0,a_1,\ldots,a_n$ (a_0 the identity or zero element) we define the sets $V_0,V_1,\ldots,V_n$ to be the sets of binary n-tuples whose scalar product with the vector $(a_1,a_2,\ldots,a_n)$ is equal to $a_0,a_1,\ldots,a_n$, respectively. Every set V_i is shown to be a single 1-error correcting code and it will be called a group theoretic code. For any Abelian group G(+), the code constituted by the vectors of V_0 is proved to have the highest information rate of all codes V_i, $i = 1,2,\ldots,n$.

Exact formulas are derived for the size of the sets V_i generated by the additive group of any Galois field. A selection criterion for choosing the best (with the highest information rate) group theoretic code of length n, i.e., determining what Abelian group G(+) of order n+1 generates the highest information rate code V_0, is presented. Many properties of the group theoretic codes are proved and several open questions of interest in the area of group theoretic codes are posed.

SOME PROBLEMS ON PERMUTATION GROUP CODES, Gérard Cohen (ENST, 46 rue Barrault 75013 Paris, France) and Michel Deza (CNRS, Paris, France). We consider mainly three directions. First, majority decoding of permutation codes. An idea of Blake, working for sharply 2-transitive groups, is generalized to some other groups, including Frobenius and sharply 3-transitive ones.

Secondly, we use permutation codes for source coding, computing distortion functions with the Hamming distance between permutations as distortion criterion, i.e., the number of elements whose rank after reproduction differs from their original rank. This problem involves weight enumeration. Finally, we list some known and new results on extremal problems linked with the minimum Hamming distance of permutation codes versus such classical parameters of permutation groups as degree, order, and degree of transitivity.

SESSION E1

CONFIGURATION MATRICES OF GROUP CODES FOR THE GAUSSIAN CHANNEL, E. Biglieri (Istituto Elettrotecnico, Università di Napoli, Italy and Department of System Science, University of California, Los Angeles) and M. Elia (Istituto Matematico, Politecnico di Torino, Italy). Given an [M,n] code for the Gaussian channel, i.e., a collection of M vectors in Euclidean n-space, its configuration matrix is the M×M matrix whose elements are the scalar products between code vectors.

The aim of this paper is to characterize the configuration matrices of group codes. The following result is proved: "Let $\underset{\sim}{C}$ be the configuration matrix of an [M,n] code. $\underset{\sim}{C}$ is the configuration matrix of a group code if and only if a) the rows of $\underset{\sim}{C}$ are permutations of the first row; and b) a matrix $\underset{\sim}{J}$, all of whose elements are unity and whose order j is not greater than (M-1)!, exists such that the matrix $\underset{\sim}{C} \otimes \underset{\sim}{J}$ ($\otimes$ the Kronecker product) commutes with all the matrices of the right regular representation of a group $\mathcal{G}$ with order $|\mathcal{G}| = jM$."

A finite procedure is also exhibited that allows one to test if a given matrix satisfies requirement (b).

AN ALGORITHM FOR CODING UNDIRECTED GRAPHS, K. N. Venkataraman (Computer Group, Tata Institute for Fundamental Research, Bombay, India) and K. Thulasiraman (Computer Centre, Indian Institute of Technology, Madras, India). This paper is concerned with the construction of a code for undirected graphs. The code, which is a string of characters, completely defines the graphs so that two graphs have the same code if and only if they are isomorphic.

The algorithm to be presented here is based on the work of Shah, Davida, and McCarthy. A counter example is given to show that their algorithm fails to produce the optimum code for certain graphs. An efficient algorithm for the construction of the optimum code is then presented. Though this algorithm was motivated by the work of Shah et al., most of the ideas used here are different from theirs.

NEW CLASSES OF SEQUENCES WITH GOOD CORRELATION PROPERTIES, Dennis A. Shedd and Dilip V. Sarwate (Coordinated Science Laboratory and Department of Electrical Engineering, University of Illinois, Urbana, Illinois 61801). A new method is presented for the design of sequences with good correlation properties. Given a pair of m-sequences of period p^n-1, we can construct another sequence of period p^n-1 with the property that the out-of-phase autocorrelation of the new sequence is zero. These sequences are, in general, multi-valued, but proper choice of the original m-sequences enables us to obtain sequences that take on values 0, +1, and -1. The new technique also enables us to construct new classes of multi-valued sequences with good cross-correlation properties,

using either the uncorrelated binary MacWilliams sequences or the Gold sequences with bounded cross-correlation. Finally, we show that given two impulse-equivalent sequences of length N, one can design impulse-equivalent sequences of length 2N-1 without resorting to Huffman's procedure.

ON LINEARIZATION OF NONLINEAR COMBINATIONS OF LINEAR SHIFT REGISTER SEQUENCES, Tore Herlestam (Bankoårdsgatan 74, S-252 60 Helsingborg, Sweden). In this paper a new and direct proof is given of a strengthened version of Hadamard's multiplication theorem in an algebraic case. The analytic functions considered are rational forms over a Galois field, representing linear shift register sequences.

Let G(f) and G(g) denote the linear spaces of all shift register sequences over the finite field $GF(p^m)$, with feedback polynomials f and g, respectively. The multiplication theorem states that

1. there esists a polynomial f∧g over $GF(p^m)$, depending on f and g only, such that $a \wedge b = (a_k b_k) \in G(f \wedge g)$ for every $a \in G(f)$, $b \in G(g)$

2. every zero of f∧g is a product of a zero of f times a zero of g,

3. if u is a zero of f and v of g, of multiplicities r and s, respectively, then uv is a zero of f∧g of multiplicity $\leq r+s-1$,

4. $\deg f \wedge g \leq \deg f \cdot \deg g$,

5. if f and g have exclusively simple zeroes and all products uv are different, then equality holds in 3. and 4.

As applications of this multiplication theorem, it is easy to correctly predict the maximally attainable linear complexity of all kinds of nonlinear combinations of linear shift register sequences.

RECENT RESULTS ON ALGEBRAIC SOFT-DECISION DECODING, T. Y. Hwang (School of Computer and Information Science, Syracuse University, Syracuse, NY 13210), N. Q. Duc (Telecom Australia, Research Laboratories, 59 Lt. Collins Street, Melbourne, Victoria, Australia 3000) and C. R. P. Hartmann and L. D. Rudolph (School of Computer and Information Science, Syracuse University, Syracuse, NY 13210). In this paper, we present the most recent results of a continuing investigation of algebraic soft-decision decoding of linear block codes using the dual code domain. It is known that any decoding algorithm for a binary code with minimum Hamming distance d which corrects all errors within an Euclidean sphere of radius $\sqrt{d}/2$ (i.e., a "maximum-radius" decoding algorithm) is asymptotically optimum for the AWGN channel. We show that this condition is not only sufficient for asymptotic optimality, but is also necessary. It is also known that any L-step orthogonalizable code can be maximum-radius

decoded using the orthogonal parity checks and the cosine demodulation function. We extend this result to include a class of hard-limited linear demodulation functions. Then the relationship between the Euclidean radius of the decoding sphere and the combinatorial structure of the parity check set is extended to nonorthogonal checks. Since algebraic soft-decision decoding using the dual code domain operates entirely in the domain of continuous functions over the complex numbers, it lends itself naturally to iterative decoding techniques. The hill-climbing approach to decoding is described and simulation results are given.

SESSION E2

OPTIMAL SAMPLING OF INDEPENDENT INCREMENT PROCESSES, B. W. Stuck (Bell Laboratories) and C. M. Newman (Indiana University). An independent increment process with one of two sets of parameters is observed. New necessary and sufficient conditions are found for the probability measures of the two processes not to be mutually orthogonal. A discrete time analog of this continuous time problem is studied, where we fix the number of observations and study Chernoff type bounds on various error probabilities as a function of the sampling interval. A partial solution is obtained to the problem of choosing the optimal sampling rate to minimize various error probabilities.

A PREDICTION PROBLEM, David Slepian (Bell Laboratories and University of Hawaii). N equally spaced samples of a bandlimited white noise are combined linearly to form an estimate of the next sample that has minimum average squared error. How does this smallest error variance decrease with N and how does it depend on the interval between samples? An asymptotic solution to the problem is obtained using properties of the discrete prolate spheroidal wave functions.

PROPERTIES AND APPLICATIONS OF A USEFUL CLASS OF TWO-DIMENSIONAL RANDOM FIELDS, R. W. Fries and J. W. Modestino (Electrical and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY 12181). A useful class of two-dimensional (2-D) random fields is described and several of their more important properties discussed. This class of random fields is modeled as a marked point process evolving according to a spatial parameter. The autocorrelation and power spectral density properties are investigated for some special cases. Several applications are discussed including the modeling of real-world imagery possessing inherent edge structure. A class of efficient edge detection algorithms is described based upon 2-D least mean-square Wiener filtering concepts where the stochastic model for edge structure is chosen from this class of random fields. Additional applications drawn from the image coding area are described.

EXTRAPOLATION OF HOMOGENEOUS RANDOM FIELDS, Hisanao Ogura (Department of Electronics, Kyoto Institute of Technology, Matsugasaki, Kyoto, 606, Japan). An extrapolation formula is given for a homogeneous random field with discrete parameters, corresponding to Helson and Lowdenslager's regularity condition related to the extrapolation in terms of the "half" space values. The process of extrapolation consists of two extrapolation processes similar to the one-dimensional prediction. The homogeneous random field is also represented in terms of the innovation field associated with the given

extrapolation problem. The extrapolation of a continuous-parameter random field can be obtained from the discrete-parameter theory just as in the prediction theory of a stationary process. Several examples of homogeneous random fields are studied using the extrapolation formula in order to examine some specific aspects, such as the Markovian property.

ENVELOPE CONSTRAINED FILTERS WITH UNCERTAIN INPUT, Robin J. Evans and Antonio Cantoni (Department of Electrical Engineering, University of Newcastle, Australia, 2308). In the envelope-constrained filtering problem with uncertain input, the filter impulse response is optimized subject to the constraint that its output lies within a prescribed envelope when the input to the filter is a signal belonging to the set defined by an input envelope constraint. The technique has application in a variety of areas such as robust radar sidelobe suppression filter design, fixed equalizer design for communication channels and system modelling with unknown but bounded noise on the input and output measurements. The results presented in the paper constitute a significant extension of the recently developed envelope-constrained filtering technique where the filter impulse response is optimized subject to constraints on the output pulse shape.

MOMENT FORMULAS FOR THE NUMBER OF SLOPE-CONSTRAINED LEVEL CROSSINGS BY A WIDE CLASS OF STOCHASTIC PROCESSES, Douglas R. Anderson and Daniel D. Carpenter (TRW DSSG, Redondo Beach, CA 90278). Level crossings constrained by upper and lower bounds on the slope at the crossings arise in such communications problems as transmission over the channels with multipath fading or random polarization. For fading channels, counting only the downcrossings of a level ℓ with slope exceeding a given value "a" produces the number of deep fades characterized by ℓ and a. Evaluation of the k-th factorial moments of the number of level crossings permits the development of a series formula for the distribution of the time between the n-th slope-constraint passage time through readily adaptable formulas for unconstrained level crossings. The authors have shown that the number of slope-constrained upcrossings (downcrossings) is a well-defined random variable with first and higher order moment formulas generalizing the unconstrained results of Rice and Cramer and Leadbetter with the following assumptions. A separable process is defined on a finite time interval with 1) a derivative that exists as a limit in probability and is continuous in probability and 2) almost surely absolutely continuous sample functions. These assumptions admit even separable processes with derivatives that exist and are continuous in the mean of any order as low as one. Let N_a^+ be the number of these upcrossings of the level ℓ by X(t)

such that X'(t) has an infimum greater than a in some neighborhood of each upcrossing. If the joint density function of X(t) and X'(t) exists and is continuous, then we have for the k-th factorial moment

$$E[N_a^+(N_a^+-1)\ldots(N_a^+-k+1)] = \int_0^1 \cdots \int_0^1 \left[\int_a^\infty \cdots \int_a^\infty y_1 \cdots y_k p_{\underline{X}(t),\underline{X}'(t)} \right.$$

$$\left. (\ell,\ldots,\ell;\ y_1,\ldots,y_k)dy_1\ldots dy_k \right] dt_1\ldots dt_k$$

where $\underline{X}(t) = (X(t_1),\ldots,X(t_k))$ and $\underline{X}'(t) = (X'(t_1),\ldots,X'(t_k))$

TREE ENCODING OF SPEECH AT 8000 BPS, Stephen G. Wilson (Department of Electrical Engineering, School of Engineering and Applied Science, University of Virginia, Charlottesville, VA 22901). Tree encoding of speech and speech-like stationary Gaussian sources is examined with an encoding rate of 8000 bps. Souce coding relative to the selected distortion measure is accomplished with the M-algorithm, which extends M survivor tree paths at each level of the tree.

The paper focuses on two extensions of previous work reported by Anderson et al. to improve performance at the lesser bit rate. A frequency-weighted error criterion is applied as an attempt at improving the subjective quality of speech. Specifically, a discrete-time weighting network approximating the response of the C-message weighting network is developed and applied to the coding of a stochastic process having long-term speech statistics. Tree searching with the M-algorithm using weighted error is shown to produce error spectra closely matching the theoretically optimal allocation, and the total weighted error is about 1.5 dB from the rate-distortion bound at M = 16. Subjective evaluation of speech coded in this fashion reveals a preference for the weighted search relative to the standard squared-error case.

Second, adaptation of the tree encoder on a short-time interval basis is considered as a means of tailoring the source code to the various quasi-stationary modes of speech. In conjunction with tree coding as above, a signal-to-distortion power ratio of approximately 15 dB was obtained for a typical sentence. A sample of processed speech at 8000 bps will be played.

TREE ENCODING OF SPEECH BASED ON SHORT TERM AUTOCORRELATIONS, S. Mattsson (Department of Electrical Engineering, Linköping University, Linköping, Sweden). Low rate source encoders for speech-like signals usually employ tree encoding with a feedback-structure, where a predictor in the feedback loop (the code generator) produces an estimate of the current input sample on the basis of a digital input from a search algorithm working on the prediction error.

All practical search algorithms (including the Viterbi algorithm and the (M,L)-algorithm) make decisions on a small number of samples (5-25). Based on the quasistationary property of speech, we investigate by simulation the performance of a binary (one bit per sample) tree encoding system, adapting to the different modes of stationarity. The code generator is chosen to have an autoregressive structure where the coefficient vector belongs to a small set of vectors, calculated on basis of the short-term autocorrelation (of a typical speaker). The search algorithm, consisting of the (M,L)-algorithm, chooses for each sequence of L speech samples that code generator (one of 2^n possible) which produces the best (minimum MSE) estimate of the actual sequence. Information (n bits) about the chosen

code is transmitted to the decoder in each block of L binary path map digits. (The set of possible coefficient vectors is assumed to be stored in the decoder.) Simulation results for different choices of code generator ensembles will be presented.

STACK ALGORITHM SPEECH ENCODING, S. Mohan and J. B. Anderson (Department of Electrical Engineering and Communications Research Laboratory, McMaster University, Hamilton, Ontario, Canada, L8S 4L7). In recent years, there has been increasing interest in the design of efficient source encoding algorithms. Source coding involves a twofold design problem, design of a good code, and design of an effective search algorithm to explore among the code words. Recent experimental research has provided good speech codes; good codes for theoretical sources are well known. A number of code search algorithms have been proposed, and we consider here the performance of one, the stack algorithm, on speech waveforms. In its simple form, the stack algorithm is not easily instrumentable, since its stack can grow very large. We thus impose constraints on the width and size of the stack. Furthermore, we divide the stack into two stacks, a main stack and an extension stack, which is periodically merged into the main stack. This allows the use of a computationally more efficient sort/merge technique. Tests on actual speech with a software-implemented algorithm show a complicated dependence on the bias term in the path metric, and show that for any fixed stack size, biases in a certain range must be used. For a fixed distortion performance, there exist unique best stack size and bias combinations, which optimize either storage, overall computation time, or nodes visited/output released. Unfortunately, these combinations are all different. This behavior matches theoretical results reported earlier for the stack-encoded binary i.i.d. source with Hamming distortion measure; we conjecture it extends to many other sources.

SOFT DECISION DEMODULATION FOR PCM ENCODED SPEECH SIGNALS, Carl-Erik Sundberg (Telecommunication Theory, University of Lund, Fack, S-220 07 Lund, Sweden). The effect of digital errors in PCM encoded speech signals transmitted over a noisy channel is reduced by using soft decision demodulation at the receiver. The reliability numbers supplied by the soft demodulator are used to point out likely transmission errors, especially in the most significant PCM bits. When a likely transmission error is identified, the corresponding PCM word is rejected by the receiver and replaced by a predictor estimate or an interpolation estimate if delayed decisions are used.

We have analyzed soft decision demodulation for standard PCM encoded speech signals transmitted over the Gaussian channel with coherent PSK (phase shift keying) or DPSK (differentially encoded PSK with coherent detection). A signal-to-noise ratio gain in E_b/N_o of the

order of 1-2 dB is obtained at low input signal levels. The gain depends on the performance of the predictor or alternatively the interpolator. No modifications of the transmitter are required to obtain this improvement. The comparisons are made with hard decision demodulation.

MATHEMATICAL TREATMENT OF SPEECH SIGNALS, D. Wolf and H. Brehm (Institut für Angewandte Physik der Universität Frankfurt am Main, Frankfurt am Main, FR Germany). Standard telephone speech signals may be considered to be realizations of a stationary and ergodic random process. The first order amplitude probability density (pd) is well fitted by the zero order modified Bessel function. Thus, the process can be represented by the product of two mutually independent gaussian random processes. Principally, this model allows us to derive higher order pd's. The mathematical treatment is considerably simplified by the assumption of elliptical invariance. This assumption has been suggested by experimental results on the statistical dependencies between amplitudes separated not more than 5 ms.

Starting with $p_1 = \pi^{-1} K_o(|z|)$ one obtains the higher order pd's

$$p_2 = M_2^{-1/2}(2\pi r_2)^{-1}\exp(-r_2),$$

$$p_3 = M_3^{-1/2}(2\pi^2 r_3)^{-1}K_1(r_3),$$

$$p_4 = M_4^{-1/2}(4\pi^2 r_4^3)^{-1}(1+r_4)\exp(-r_4) \text{ etc.},$$

$$r_n = [\underline{z}^T \underline{\underline{M}} \underline{z}]^{1/2}, \ \underline{z}^T = (z_1, \ldots, z_n), \ \underline{\underline{M}} = (\langle \zeta_i \zeta_k \rangle), \ M_n = \det \underline{\underline{M}}.$$

These results offer a key to an analytical treatment and the solution of many practical problems in the field of digital processing of speech. For example the effect of linear prediction and gain control on the first order pd will be discussed in more detail.

RECOGNITION OF WORDS FROM A REGULAR LANGUAGE IN THE PRESENCE OF NOISE, R. L. Kashyap (School of Electrical Engineering, Purdue University, West Lafayette, IN 47907). Consider an alphabet $A = \{a_1, \ldots, a_m\}$ made of m distinct symbols and a set $B = \{x_1, \ldots, x_{m_1}\}$ where each word x_i is a concatenation of symbols from A with

no adjacent symbols identical. Suppose we are given a string z, a concatenation of symbols in A in which adjacent symbols could be the same. z is a noise corrupted version of some $x \in B$, where the noise can delete some symbols of x or insert some symbols into x or replace some symbols of x by other symbols. Our problem is to find the word in B which is "nearest" to the string z according to a suitable distance measure D(z,x). The available distance measures between strings are not appropriate for this problem in view of the special condition on the members of B. We define a new distance measure between z and x in terms of the 2 operations of deletion and replacement (no insertion) operations needed to convert z into x, the weights attached to the operation depending on the symbol involved. The weights are chosen so that D(z,x) can be interpreted as the negative of the likelihood function that z could have been generated by x.

This distance measure is particularly appropriate in some speech recognition problems. We given an algorithm to determine the x which minimizes the value D(z,x). Unlike some of the algorithms in the literature, we need to compute D(z,x) only for a small number of x belonging to B.

ESTIMATION OF PHONEMIC CONFUSION USING THE CONCEPT OF MAHALANOBIS DISTANCE, S. Tazaki, H. Osawa, Y. Yamada and T. Gotoda (Ehime University, Matsuyama, 790, Japan). The problem of estimating quantitatively the possibility of phonemic confusion is described. This problem originates from the fact, which has already been reported in various journals, that the fluctuation of cues which results from differences in following vowels and differences in speakers affects listeners' ability to recognize and distinguish among phonemes. Further we have discovered that we need to make use of several cues in the frequency domain as well as in the time domain for doing reliable discrimination of phonemes. In the first place, the concept of Mahalanobis distance, one of the Multivariate Analysis methods, is applied to set up an efficient metric between two different isolated syllables because this distance is easily calculated by using a 'calculus of variation' so as to minimize the variance within each syllable and maximize the variance between the two syllables. Next, the syllables which belong to the same phoneme are arranged systematically.

We regard the deviation from the average metric value to be the range of fluctuation in a phoneme. We also regard the average distance between the different phonemes to be the range of interference with each other phoneme. As a result, in the case of Japanese stop consonants, it was found that over 80 percent of our estimates could be shown to correspond with actual data, and in addition, for the first time to our knowledge the direction of confusion could be estimated, at the rate of about 80 percent.

SESSION E4

SHANNON'S "PROOF" OF THE NOISY CODING THEOREM, James L. Massey (Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139). It is often said, particularly by mathematicians, that Shannon nevery really "proved" his Noisy Coding Theorem which asserted that the decoding error probability can be made arbitrarily small at any code rate R less than channel capacity C. It will be shown on the contrary that Shannon's original "proof," in the case of a discrete memoryless source (DMS) and a discrete memoryless channel (DMC), can be made completely rigorous using only methods well within the grasp of an average undergraduate engineering student. It will be argued that the dissatisfaction with Shannon's "proof" is more likely traceable to the failure to see that Shannon's proof considers two quite distinct situations [namely, (1) a "gedanken" experiment to determine the good candidates for code words, and (2) a coding scheme for transmitting the DMS through the DMC], than to the "incompleteness" of Shannon's "proof."

A NEW LOOK AT EXPONENTIAL ERROR BOUNDS FOR MEMORYLESS CHANNELS, I. Csiszár, J. Körner, and K. Marton (Mathematical Institute of the Hungarian Academy of Sciences, H-1053 Budapest, Reáltanoda u.13-15, Hungary). We prove a non-probabilistic lemma on the existence of a "balanced" subset of $\mathcal{X}^n$ of size $\exp(nR)$. With this codeword set, the mutual information distance decoding rule achieves Gallager's and Forney's exponential error bounds universally for DMC's with input alphabet $\mathcal{X}$ when restricting attention to fixed composition codes. The main interest of these results is that both the codeword set and the decoding rule are independent of the channel. This means, in particular, that for the class of DMC's with critical rate for the given composition less than R, there exist asymptotically optimal universal codes of rate R.

Further applications of the lemma as well as new bounds on the probability of decoding error on a DMC with noiseless feedback will be discussed.

RELIABILITY FUNCTION OF A DMC AT RATES ABOVE CAPACITY, J. Körner (Mathematical Institute, Hungarian Academy of Sciences). It is known as Wolfowitz's strong converse to the coding theorem for a discrete memoryless channel (DMC) that at rates above capacity the probability c(R,n) of correct decoding of the "best" n-length block code of rate at least R tends to zero. Here,

$$c(R,n) \triangleq \max_{\mathcal{C}} \min_{\underline{x}\in\mathcal{C}} c_{\underline{x}}$$

where the maximum is taken over all codes $\mathcal{C}$ of rate at least R and block-length n, and $c_{\underline{x}}$ is the probability that no error occurs when the codeword $\underline{x}$ of the code $\mathcal{C}$ is transmitted. Arimoto has given an exponential upper bound on c(R,n) having a form similar to Gallager's expression of the random-coding exponent.

In this paper we give a simpler derivation of the same bound in the now familiar divergence extremum form and show that the bound is tight at all rates above capacity. The result is that

$$\lim_{n\to\infty} -\frac{1}{n} \log c(R,n) = \min_{P} \min_{\hat{W}} \sum_{x\in X} P(x) D(\hat{W}(\cdot|x)||W(\cdot|x)) + [R-I(P,\hat{W})]^{+}$$

where the first minimum refers to all probability distributions P on the input alphabet X of the DMC $W:X\to Y$, and the second minimum is for all "test channels" $\hat{W}:X\to Y$. $D(\hat{Q}||Q)$ is the informational divergence (or discrimination) and $[t]^{+} \triangleq \max [t,0]$.

This result has theoretical interest inasmuch as it shows that for R>C there is no critical rate, i.e., no "break in the behavior of the average code" occurs.

BLOCK CODING FOR DISCRETE STATIONARY $\bar{d}$-CONTINUOUS NOISY CHANNELS, Robert M. Gray and Donald S. Ornstein (Electrical Engineering Department and Mathematics Department, Stanford University, Stanford, CA 94305). A new class of discrete stationary noisy channels with memory termed $\bar{d}$-continuous channels is introduced and shown to include all stationary discrete channels for which coding theorems exist. Roughly speaking, a $\bar{d}$-continuous channel is such that the effect of the "past" and "future" inputs on the "present" outputs dies out asymptotically as measured in a $\bar{d}$ or average Hamming distance sense. This is weaker than the corresponding notions of Pfaffelhuber, Kadota, and Wyner, who require that probabilities of all n-tuples be close; that is, closeness in a variational or distribution sense.

General block channel coding and block joint source and channel coding theorems are proved for stationary $\bar{d}$-continuous channels, and various definitions of channel capacity are compared.

RANDOMNESS IN DISCRETE CHANNELS WITH MEMORY, P. C. Shields (University of Toledo, Toledo, OH 43606) and D. L. Neuhoff (University of Michigan, Ann Arbor, MI 48109). A result on the nature of the randomness in discrete channels with memory is presented. It is shown that certain channels, including finite-memory channels, may be decomposed into a memoryless channel followed by a filter. Equivalently, it is shown that such channels may be simulated with an external, memoryless source of randomness and a sliding-block code. The channels considered here are $\bar{d}$-continuous (in the sense of Gray and Ornstein) and conditionally almost block independent (a notion related to very weak Bernoulli random processes). Channels in this class have almost finite input memory and almost finite output memory. A measure of channel similarity called the $\bar{d}$-channel distance is introduced and it is shown that the above channels are precisely the class of channels which can be well approximated by finite-memory channels. We conclude that the randomness in such channels is in no way different than the randomness in a memoryless channel or memoryless source and channels not of this type cannot be well-modelled by finite-memory channels.

METRICS MATCHED TO THE DISCRETE MEMORYLESS CHANNEL, Gérald Séguin (Department of Electrical Engineering, Royal Military College, Kingston, Ontario, Canada). A metric D is said to be matched to a discrete memoryless channel (DMC) if it is the sum on its coordinates of a single letter metric and if the maximum likelihood decoder is a minimum D-distance decoder. Necessary and sufficient conditions on the transition probabilities of a DMC for the existence of a matched metric are derived. In the case of the binary DMC these are equivalent to the channel being symmetric. Explicit transition probabilities are given for a large class of ternary DMC with an associated matched metric.

THE GEOMETRY OF PROBABILITY SPACE, Richard E. Blahut (International Business Machines, Inc., Owego, NY 13827). The space of all probability distributions on a finite set of size K elements is made into a vector space by suitable definitions of vector addition and scalar multiplication. Distance and size measures are studied within the setting of this vector space. This structure yields bounds on the performance of testers of binary or multiple hypotheses and of channel codes.

SESSION E5

BOUNDS ON UPDATE ALGORITHMS, Richard A. Flower (Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801). Updates in a computer-implemented data system are investigated using a data model developed by Elias and Welch. In the model a data base is recorded by storing a representative binary string in computer memory. A retrieval question is answered by reading individually addressable memory cells, each of which stores a single binary value of the representative string. An update to a new data base is made by an algorithm which reads and/or writes at some memory cells and halts when the string in memory is one of the strings representing the new data base.

Memory costs and retrieval costs are two characterizations of system performance. These costs have been investigated by Elias. Here we investigate update costs, which are another important characterization of system performance.

STOCHASTIC ERROR ANALYSIS OF SPLINE APPROXIMATION, H. L. Weinert (Electrical Engineering Department, Johns Hopkins University, Baltimore, MD 21218), G. S. Sidhu (Instituto de Ingenieria, University Nacional de Mexico, Mexico 20, D.F.) and R. H. Byrd (Math. Sciences Department, Johns Hopkins University, Baltimore, MD 21218). In order to use traditional spline approximation error bounds, one needs at the very least a tight upper bound on some nonlinear functional of the unknown function producing the data. In most practical problems, however, this information is not available, and thus these bounds cannot be computed. The most one can do in this situation is bound a normalized error. This computable upper bound is in fact the mean-square error of an associated least-squares estimation problem whose statistics are determines by the type of spline used. The bound is independent of the data and can thus be used to develop optimal sampling schemes.

FITTING CURVES TO DETERMINISTIC DATA, H. L. Weinert and R. H. Byrd (Johns Hopkins University, Baltimore, MD 21218) and G. S. Sidhu (Instituto de Ingenieria, University Nacional de Mexico, Mexico 20, D.F.). Many of the optimal curve fitting and approximation problems arising in numerical analysis have the same structure as certain estimation problems involving random processes. This structural correspondence, whose development depends on results from the theories of reproducing kernels, linear systems, and stochastic realization, has many useful consequences. For one thing, it tells us that optimal curve fits (splines) are sample functions of linear least-squares estimates. As a result, recursive estimation techniques can be used to solve curve fitting problems and to provide tight upper bounds for fit error. In addition, the correspondence yields

interesting information on the sample function properties of least-squares estimates of certain random processes. Previous work has dealt with splines determined by differential operators; these splines are sample functions of estimates of autoregressive processes (generated by all-pole systems). The present work examines splines determined by certain integro-differential operators; these splines are sample functions of autoregressive-moving average processes (generated by systems with zeros).

SQUARE-ROOT ALGORITHMS FOR PARALLEL PROCESSING IN OPTIMAL ESTIMATION, M. Morf, J. Dobbins, B. Friedlander and T. Kailath (Information Systems Laboratory, Stanford University, Stanford, CA 94305). We give explicit algorithms in square-root form which allow measurements for the standard state estimation problem to be processed in a highly parallel fashion with little communication between processors. After this preliminary processing, blocks of measurements may be incorporated into state estimates with essentially the same computation that usually accompanies the incorporation of a single measurement. Other results of this formulation of the estimation problem include square root doubling formulas for calculating steady state covariance with constant models and an extension of the class of problems for which Chandrasekhar-type algorithms offer computation reductions to include piecewise constant systems with arbitrary initial conditions.

LADDER FORMS FOR ESTIMATION AND DETECTION, M. Morf, D. T. Lee and A. Vieira (Information Systems Laboratory, Stanford University, Stanford, CA 94305). Many efficient algorithms in estimation theory have been based on certain canonical forms. For instance, the observer canonical form is one that has been widely used in algorithms involving state estimation via Riccati and Chandrasekhar-type equations. One class of canonical forms that provide the most promising alternative are ladder (or lattice) canonical forms. These ladder forms have the lowest computation and storage requirements and their nice property of stability and minimum-phase can be checked by inspection of the so-called "reflection coefficients". These coefficients can be identified with the partial correlation (PARCOR) coefficients from the stochastic point of view, and they also have physical significance related to the scattering theory of waves.

It was shown, for example by Itakura and Wakita, that the ladder coefficients of an autoregressive (AR) model can be computed via Levinson's recursions involving the forward and backward predictors (a(z) and b(z)). The ladder coefficients are in one-to-one correspondence with the correlation sequence, as well as to the a(z) and b(z) polynomials. Therefore it should be possible to compute the ladder coefficients without the explicit use of the predictor coefficients. The fact being that the computation for the ladder

coefficients involves an inner product of an input data sequence, e.g. the correlation sequence, with the impulse response of the predictor, a(z), it becomes clear that the ladder coefficients are sufficient to parameterize the predictor and therefore the algorithms to determine the ladder coefficient do not necessarily require the explicit use of a(z).

We show here how these algorithms can be obtained as a natural extension of the Fast Cholesky or Chandrasekhar equations similar to their derivations as an extension of Levinson's algorithms.

Then we show how to extend these algorithms to the moving-average (MA) processes via the use of feedback ladder forms. Especially for these feedback forms, the property of stability by inspection is very useful. We also consider autoregressive-moving-average (ARMA) models by combining these two results.

All these algorithms can be generalized to multichannel, nonstationary and multi-dimensional cases. We shall discuss some interesting applications of these ladder forms to problems in estimation and detection.

ON SOME COMPLEXITY PROBLEMS IN COMMUNICATION THEORY, K. Yao (Department of System Science, University of California, Los Angeles, CA 90024). Advances in microelectronics have made the study of efficient algorithms for digital processing a topic of interest. A survey of known results is made on minimal multiplications needed in performing inner product of vectors and matrix-vector multiplications with and without using preconditioning requirements. New results involving minimal multiplication realizations are obtained for several detection and extended generalized Wiener filter problems. In many cases, there can be considerable reduction in the number of multiplications with only small decrease of performance or increase of programming cost.

ON THE USE OF ASSOCIATIVE MEMORIES FOR PATTERN RECOGNITION AND INFORMATION PROCESSING, Yeh-Han Pao and W. L. Schultz (Case Western Reserve University, Cleveland, OH 44106). The characteristics of associative or content addressable memories which may be implemented with parallel processing techniques and offer efficient memory space utilization are discussed in the context of pattern recognition and adaptive control applications.

SESSION F1

MAXIMUM POSTERIOR PROBABILITY DEMODULATION OF ANGLE-MODULATED SIGNALS, D. W. Tufts (Electrical Engineering Department, University of Rhode Island, Kingston, RI 02881) and J. T. Francis (Naval Underwater Systems Center, Newport, RI 02840). Certain results of functional analysis are used to solve the problem of Maximum Posterior Probability (MPP) demodulation of angle-modulated signals. A sampled data version of the problem is considered, in which additive Gaussian noise and a Gaussian modulating signal are assumed. The results of this paper can be considered as a numerical method for solving the nonlinear integral equation which specifies the necessary condition for MPP estimation for the corresponding continuous-time problem.

The results can be applied to the generation of a partially coherent reference for improving signal detection.

PRE-PROCESSING OF SIGNALS FOR A CLASS OF CDM PARALLEL DATA TRANSMISSION SYSTEMS, F. Ghani (Department of Electrical Engineering, Aligarh University, Aligarh 202001, India). The paper presents a technique for the processing of signals at the transmitter for a class of CDM parallel systems suitable for the transmission of digital data over time-invariant or slowly time-varying channels. The processing at the transmitter is such that it maximizes output signal-to-noise ratio and no detection process is needed at the receiver. It is shown that the performance of the scheme in the presence of white Gaussian noise is similar to that of a serial arrangement where all the processing is achieved at the receiver. The proposed technique combines the advantages of FDM parallel systems and serial systems used in data transmission.

ERROR PROBABILITY FOR ZERO-CROSSING DETECTION OF DIGITAL FM, Hüseyin Abut (Bogazici University, Istanbul, Turkey). An upper bound is presented for evaluating the error probability of a binary FM system subject to additive Gaussian noise (AGN). We consider in this work a binary FM system that employs a continuous-phase FSK modulator and a zero-crossing detector. The zero-crossing (ZC) detector counts the number of zero-crossings of the signal imbedded in noise. If the count exceeds a threshold N_c then a "space" is announced; otherwise, the decision is a "mark".

An erroneous detection results due to the presence of noise either when an extra ZC occurs or the counter misses a ZC. There are four equally probable cases leading to detection errors, one of them being that of having no ZC when a mark is transmitted and the instantaneous phase of the sinusoidal process indicates that a ZC should have been present. Upper bounds to the error probability

are derived for all these cases. However, expressions for the bounds are found to be intractable and the resulting triple integrals have been computed numerically. The obtained upper bounds are remarkably good. In fact, if the baud rate is 1200 bits/sec and the mark and space frequencies are 1300 and 2100 Hz, respectively, then the upper bound on error probability is 2×10^{-5} for a signal-to-noise ratio (SNR) of 12 dB. This value is lower by a factor of ten in comparison with any other results available in the data communications literature.

OPTIMUM PCM CODES FOR DPSK, Carl-Erik Sundberg (Telecommunication Theory, University of Lund, Fack, S-220 07 Lund, Sweden). The effect of digital errors in PCM systems for speech signals depends on the PCM code used. The standard binary folded PCM code is superior to the natural binary code but it is not optimum. This paper considers PCM codes especially designed to be insensitive to the error patterns produced by differentially encoded phase shift keying (DPSK). It is concluded that the dynamic range with specially designed PCM codes is considerably extended compared to standard binary folded PCM. Special simplified interleaving schemes with optimum PCM codes for coherent PSK are also considered. The dynamic behavior of various PCM codes is calculated and compared, e.g., asymptotically optimum PCM codes for DPSK, binary folded PCM with and without interleaving, asymptotically optimum PCM codes for PSK (coherent), etc. We conclude that the best scheme for DPSK is an optimum PCM code for PSK and simplified interleaving. The reencoding from standard binary folded to the new PCM codes can, for example, be implemented by means of a read only memory.

COSYDAI - A NEW INTEGRATED COMMUNICATION SYSTEM WITH DATA COMPRESSION AND ERROR CONTROL, M. Barducci, G. Benelli, V. Cappellini, and E. Del Re (Instituto di Elettronica, Facolta' di Ingegneria, Universita di Firenze, Italy). Communication systems using both data compression and error control coding are considered. A software package to simulate systems using the two operations in cascade is described. A new integrated compression-coding strategy, called COSYDAI (COmpression with SYnchronization control and DAta Interpolation), is proposed, in which a suitable synchronization is added to compressed data and special controls are performed on the received data for error correction. By examining the received synchronization words, decisions are taken regarding the errors introduced by the communication channel and the words estimated with errors are reconstructed through interpolation operations. Results obtained by using the above simulation package, suitably modified with prediction algorithms, show how in several channel noise conditions (as with burst noise) the COSYDAI system has higher efficiency than the more standard systems which perform the operations of data compression and error correction in cascade.

ERROR RATE PERFORMANCE OF A FADING MULTICHANNEL SYSTEM, Howard E. Nichols (GTE Sylvania, Electronic Systems Group, 77 "A" Street, Needham Heights, MA 02194). This paper considers the problem of digital communication over a channel whose performance is limited by fading and interchannel interference. Using the optimum maximum likelihood detector, bit error rate performance is calculated for flat Rayleigh fading channels corrupted by interchannel interference and additive white Gaussian noise. It is shown that the performance for an N-channel system is similar in form to the performance of an N-diversity system using maximal-ratio combining.

TRELLIS CODING WITH EXPANDED CHANNEL-SIGNAL SETS, G. Ungerboeck (IBM Research Laboratory, CH-8803 Rueschlikon, Switzerland). In this paper further progress in trellis coding with expanded sets of multilevel/phase channel signals is reported. This coding approach aims at improving error performance of data links without sacrificing data rate or requiring more bandwidth or power, by encoding information at given rate with a redundant number of channel signals. Assuming soft ML decoding in the receiver, the objective in designing codes is maximizing free Euclidean distance between code sequences of multilevel/phase channel signals. First heuristically designed codes with coding gains of 3-4 dB have been presented at the last Symposium. Now the development of more powerful codes is reported which can be viewed as binary convolutional codes of rate $R=m/(m+1)$, where m is the number of bits to be transmitted per baud, followed by mapping $m+1$ coded bits into an expanded set of 2^{m+1} channel signals. For a specific mapping, termed mapping by set partitioning, a new measure for distance between binary code sequences was defined, which efficiently lowerbounds the Euclidean distance between the corresponding channel-signal sequences. The lowest bound is always achieved between some channel-signal sequences and hence equals the free Euclidean distance between these sequences. Based on the new distance measure, a search algorithm for binary convolutional codes was developed, by which, in conjunction with the mapping rule, codes with coding gains up to 6 dB were found for a wide range of practical multilevel/phase modulation forms.

SOME STATISTICAL PROPERTIES OF THE MONOPULSE RATIO, Irving Kanter (Raytheon Company, Bedford, MA). The monopulse ratio is treated under conditions of multiple non-fluctuating targets and correlated Gaussian interference between the difference and sum channels. Three theorems concerning respectively the mean, the variance and the first absolute central moment are derived. The theorems are valid for arbitrary signal-to-interference ratio; their proofs do not require knowledge of the probability density function (p.d.f) of the monopulse ratio. The results are compared to the approximate procedures currently employed by radar engineers when the signal-to-interference ratio is large or when the interference is caused by jamming signals. When the signal-to-interference ratio vanishes, the results present the mean and "spread" of the p.d.f. of the monopulse ratio when the radar confronts multiple jammers.

The appendix corrects a Soviet publication which gives the p.d.f. of the monopulse ratio as an infinite series of hypergeometric functions. It also presents a closed form solution by the author which involves only Bessel functions of order zero and one. When the first absolute central moment is calculated using each of these representations, theorem III is verified.

DETECTION AND PARAMETER ESTIMATION OF CLOSELY SPACED MULTIPLE TARGETS, U. Nickel (Forschungsinstitut für Funk und Mathematik, 5307 Wachtberg-Werthhoven, F.R. Germany). This paper is concerned with the directional resolution of M radiating sources, being not distinguishable by conventional Fourier or monopulse processing. No a priori information about the sources is assumed to be available.

The estimation portion of the problem is carried out by maximum likelihood estimation. The resulting non-linear minimization problem is investigated and a strong dependency of the conditioning on the sources' phase, amplitude and position is found. Therefore, a stochastic approximation algorithm for the minimum search seems more appropriate than a minimization based on a single observation.

The hypothesis testing portion of the problem can be solved by a sequence of one-sided tests for $M = 1 \ldots M_{max}$. For each M a decision can be made either by using knowledge of the accuracy of the estimation problem at known SNR or by using the likelihood ratio when the noise level is known.

OPTIMAL DETECTION OF A TWO-STATE MARKOV PROCESS IN NOISE, Erdal Panayirci (Technical University of Istanbul, Istanbul, Turkey). In this paper, the problem of detecting a two-state, continuous-time Markov process with a fixed observation time in additive white Gaussian noise is investigated. It is considered in both the discrete and continuous domains, and computationally feasible algorithms are derived which optimally solve the problem.

The optimum detector is defined by the Bayes decision rule and its basic components are generated iteratively. It is shown that the optimal decision rule is an element of a class of on-line decision rules with fixed memory. However, the number of computations needed to make the optimum decision grows linearly with length of the observation sequence.

In order to analyze the quality of the optimal detector, it is necessary to consider the probabilities of errors of the first and second kind. For this, the conditional error probabilities are first studied and iterative relationships established for them from which the usual probabilities are found as a function of the observation time and the parameters of the problem.

Finally, the case of continuous observation times is considered. Stochastic differential equations are obtained for the logarithm of the likelihood ratio and for the conditional probabilities of errors. Since these equations are non-linear as well as stochastic, one must not expect to obtain any exact analytical solutions for them. However, some approximate methods are given at the end of the paper.

DIGITAL SIGNAL DETECTION *WITH QUANTIZED* OBSERVATIONS, Sam Reisenfeld (Department of Systems Science, University of California, Los Angeles, Los Angeles, CA 90024). The problem considered is a binary, base-band digital communications system. The continuous time additive white Gaussian noise channel model is assumed. The channel output is low pass filtered, sampled, quantized and put through a digital signal processor. The digital signal processor output is the detected bit stream. L time samples are used in each bit decision, and the quantizer has N levels. The digital signal processor is ideal in the sense that it has no internal quantization error. The quantizer preceding the digital signal processor imposes a fundamental limitation on the detection performance, and this limitation is investigated.

The performance measure which is used is P_e, the probability of decision error. The P_e performance is investigated as a function of E_b/N_o, L, and N. For N = 2, or binary quantization, it is shown that,

$$\underset{1 \le L < \infty}{\text{l. u. b.}} \; P_e = \lim_{L \to \infty} P_e = \operatorname{erfc}\left(\left[\frac{2E_b}{N_o} \frac{2}{\pi} \right]^{1/2} \right).$$

Therefore, binary quantization of the channel output implies a maximum performance loss of $\pi/2$ or 1.961198771 dB.

For N greater than 2, the joint distribution of the number of times each quantizer output is reached in one bit interval is multinomial. The optimal signal detection algorithm is derived for all L and N. The special case asymptotic form of this algorithm as L or N becomes large is investigated. The asymptotic P_e performance of the optimal detector as L becomes large is also investigated.

DIGITAL DETECTION OF PERIODIC SIGNALS IN GAUSSIAN NOISE, Mario A. Blanco (Electrical Engineering Department, Illinois Institute of Technology, Chicago, IL 60616). The problem of digital detection of periodic signals with random phase in additive gaussian noise is considered. An interesting class of digital receivers for the detection of these signals based on hardlimiting, sampling and correlation operations is presented. The optimization of these receivers is considered and it is shown how one can approximate the operation of the maximum likelihood detector by very simple correlation operations. An analysis of the statistical performance of these digital detectors is then presented, and it is shown that under some simplifying assumptions one can obtain closed form expressions for the expected value, variance and distribution function of the output of the detector. Some of these results are then extended to the case in which the signal to be detected has random phase, and random time of arrival. Also, some examples of the performance of these detectors as a function of signal-to-noise ratio and length of observation interval are presented.

DETECTION OF WEAK SIGNALS IN NARROWBAND NON-GAUSSIAN NOISE, J. W. Modestino and A. Ningo (Electrical and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY 12181). A class of nonlinear receiver structures is described for the detection of weak signals in non-Gaussian narrowband noise. In particular, the concept of a locally optimum receiver structure is extended to the case of narrowband signal and noise models. A useful class of non-Gaussian narrowband noise models is developed for which the locally optimum receiver implementation is explicitly determined. These structures are shown to provide considerable improvement over conventional linear receiver structures. The basis of comparison is taken as the asymptotic relative efficiency (A.R.E.). Unfortunately, the locally optimum receiver requires explicit a priori knowledge of the underlying noise distribution. To circumvent this difficulty a rather simple adaptive nonlinear receiver structure is described which attempts to adapt to the unknown prevailing noise environment. This adaptive receiver is shown to provide fairly efficient and robust performance in a wide variety of non-Gaussian narrowband noise environments.

SESSION F2

LEADING EDGE ESTIMATION ERRORS, Israel Bar-David and David Anaton (Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa, Israel). The leading-edge-estimator (LEE) of the time-of-arrival (TOA) of a pulse signal is defined as the instant at which the noise-contaminated and filtered version of the received signal crosses a preset threshold level. A rigorous analysis, valid for rectangular pulses in additive white Gaussian noise, indicates that the probability density of the LEE is asymmetrical and therefore the LEE is biased. Furthermore, for a pre-specified probability of anomalous estimation error, the minimum bias is inversely proportional to the signal-to-noise ratio R while the minimum variance is inversely proportional to R^2. Results are presented in the form of parameterized graphs.

MULTI-MODAL ESTIMATORS OF PULSE PARAMETERS FOR THRESHOLD-EXTENSION DEMODULATORS, I. Bar-David and B. G. Goldberg (Faculty of Electrical Engineering, Technion - Israel Institute of Technology, Haifa, Israel). The multi-modal (MM) estimator of a non-linearly involved signal parameter is a generalization of the maximum likelihood (ML) estimator in that it records M relative maxima, and the corresponding modes of the likelihood function associated with the parameter. The additional modes contain valuable information about the parameter in cases when anomalous noise-induced maxima override the signal-induced one. A class of ad-hoc algorithms is introduced that exploits this information as well as the assumed smoothness of the message that underlies the sequence of transmitted parameters, in order to lower the signal-to-noise ratio at which the noise-threshold effect occurs. Monte-Carlo simulation of a typical algorithm applied to a 2-modal estimator of speech signal exhibits a 2 to 3 dB improvement in demodulating speech and R.C. filtered noises, as expected from a plausibility argument.

SESSION F3

EXTENSION OF AN ADAPTIVE DISTRIBUTED ROUTING ALGORITHM TO MIXED MEDIA NETWORKS, Anthony Ephremides (Electrical Engineering Department, University of Maryland, College Park, MD). By modeling the broadcast portion of a mixed media network as a fully connected point-to-point network with link capacities varying as functions of the traffic rate, it is possible to extend an adaptive distributed routing algorithm that was originally developed for point-to-point ground networks. Additional modifications for improved dynamic performance at the satellite interface message processors are also included.

ANALYSIS OF AN ADAPTIVE ROUTING STRATEGY FOR COMPUTER COMMUNICATION NETWORKS, Tak-Shing Yum and Mischa Schwartz (Columbia University, New York, NY). Adaptive routing strategies have been adopted for such store-and-forward data networks as the US ARPA network and the Swedish TIDAS network. They have been studied, primarily with the use of simulation, by many investigators. Analysis of adaptive routing algorithms for networks, involving the interactive operation of networks of queues, has not received as much attention because of its complexity.

In this paper we present an approximate analysis of an adaptive routing strategy that is related to a strategy studied, using simulation, by the British National Physical Laboratory. The strategy consists of sending messages via the shortest path route if it is unique. If two or more alternate paths with the same number of links to the destination node exist, the one with the shortest outgoing queue at the current link is chosen. The analysis at any one node involves a variant of the "Join the Shortest Queue" problem and has been carried out using a Markov chain finite buffer approach. (The variant is due to the fact that the traffic into any queue has both a fixed component, and a variable component due to the shortest path algorithm.)

Simulation has indicated that the interdeparture times for messages leaving queues are closely exponential. This makes queues at different nodes effectively independent, except for a set of parameters coupling the flows together to ensure flow conservation holds. These parameters were found using a relaxation scheme.

The analysis described has been applied to an example four-node network. Various traffic matrices were assumed, both symmetrical and asymmetrical. In all cases this adaptive routing strategy resulted in a substantially lower average message time delay than the best fixed strategy minimizing average time delay. The analytic results were found to agree with simulation results to within 2%.

OPTIMAL ROUTING FOR LINE-SWITCHED DATA NETWORKS USING DISTRIBUTED COMPUTATION, Adrian Segall (Department of Electrical Engineering, Technion - Israel Institute of Technology, Haifa, Israel). An algorithm for optimal routing in line-switched data-communication networks is presented. The algorithm is such that every node in the network uses the information received from its neighbors to update its information and to reroute messages. The algorithm has the property of producing at every instant loop-free routing, and for stationary inputs and fixed topology, it reduces the average delay in the network at every step and converges to the optimal routing. In addition, it has the important property of not starting a new updating step before the previous one is completed.

PERFORMANCE OF QUEUEING SYSTEMS, M. K. Nguyen (RCA, Government Communication System, Camden, NJ 08102) and R. L. Pickholtz (George Washington University, Washington, DC). Queueing problems are central to various aspects of data networks. Many efforts have been spent on the theory of queues in finding explicit analytical solutions. These solutions are often too complicated to be practical. The present paper is intended to give some inequalities for various quantities of interest in queueing system performance. The inequalities for queue length distribution in GI/G/1 have been established (in both transient and equilibrium states). An application example for the expected queue size in a bulk queue has also been derived. Finally the inequalities of expected queue size are given in the case of stations having finite K waiting places.

ON A STOCHASTIC INTEGRAL EQUATION MODEL FOR MARKOV QUEUEING NETWORKS, Frederick J. Beutler (Computer, Information and Control Engineering Program, The University of Michigan, Ann Arbor, MI 48109). A set of stochastic integral equations is presented as an axiomatic formulation for a class of queueing networks such as might be encountered in communication or computer systems. Certain requirements on arrival and departure functions assure that the vector of queue lengths at the server nodes is well-defined as a Markov process, which behaves in accordance with the usual notion of a queueing network. At the same time, the model possesses the flexibility to represent queueing parameters such as blocking, jockeying, customer losses, variable service rates, and multiple servers at each node.

Necessity conditions for equilibrium are derived in terms of the integral equations; for Jackson networks, these coincide with the traffic equation. The infinitesmal generator for the Markov process is related to statistical equilibrium and the balance equation. The infinitesmal generator is then exhibited in explicit form.

The set of stochastic integral equations can be augmented to account for customer streams. The augmented process is also Markov, and its infinitesmal generator can be obtained with little difficulty.

THEORETICAL UPPER BOUNDS ON SPECTRAL-SPATIAL UTILIZATION IN A CELLULAR LAND MOBILE COMMUNICATIONS SYSTEM, G. R. Cooper and R. W. Nettleton (School of Electrical Engineering, Purdue University, West Lafayette, IN 47907). This short paper presents theoretical upper bounds on spectral-spatial utilization in a cellular land mobile communications system.

Spectral-spatial utilization (SSU) is defined for the purposes of this paper to be the maximum number of similar message channels per MHz for each cell of the system. It is assumed that an analog message such as speech is being transmitted on each channel, and message quality is held constant as a parametric criterion.

Two possible spectral configurations are considered:

1. The use of orthogonal channels, e.g., frequency-division multiplex, in which channels are re-used on several cells separated by a suitable geographic distance.

2. The use of non orthogonal channels, e.g., spread-spectrum with code-divi ion multiplex, in which each channel is permanently assigned to a single user.

Both these configurations are interference-limited in performance: the first by a relatively small number of co-channel sources distributed in the service area, and the second by large numbers of sources uniformly distributed over all the service area.

Under simple but not unreasonable assumptions on the nature of the service area, it is found that the bound on the SSU of a non-orthogonal-channel system is always significantly higher than the bound on the SSU of the orthogonal-channel system, over a wide range of message quality criteria and propagation characteristics.

The authors are currently investigating a spread-spectrum implementation of the non-orthogonal channel concept.

A PROTOCOL FOR RESOLVING CONFLICTS ON ALOHA CHANNELS, John Capetanakis (22 Blake Street, Cambridge, MA 02140). When a number of sources have independent access to a common channel, conflicts normally arise when more than one source attempt to access the channel simultaneously. Two algorithms have been proposed to resolve these conflicts, the TDMA and the Aloha. The problem with these protocols, however, is that they suffer either from congestion instabilities or from low throughput and high mean delay.

We have developed a class of multiple access tree search algorithms that do not have the above difficulties. We will present these algorithms and show that they are stable, have a throughput of .43 packets/slot, and have superior delay characteristics. Furthermore, we will discuss their optimality and show that TDMA is a special case of the tree algorithms.

INTEGRATED RANDOM-ACCESS RESERVATION SCHEMES FOR MULTI-ACCESS COMMUNICATION CHANNELS, Izhak Rubin (Department of System Science, School of Engineering and Applied Science, University of California, Los Angeles, CA 90024). Dynamic demand-assignment schemes, governing the access-control discipline for a network of terminals communicating through a multi-access communications channel, are presented. A repeater is employed to allow a fully-connected network structure. The channel can be characterized as inducing a low propagation-delay value, as for terrestrial radio or line networks; or as being associated with a higher propagation-delay value, as for a satellite communication channel. A synchronized (slotted) communication medium is considered.

Recent results we have derived for the delay-throughput performance of dynamic reservation and group random-access access-control schemes are presented. Integrated random-access reservation schemes are then presented and studied. These schemes adapt their protocol to the underlying traffic load using combined random-access and reservation access-control procedures. Such procedures are subsequently shown to yield excellent delay-throughput performance characteristics over the whole range of network traffic intensity values.

MULTIPLICATIVE MULTIPLE-ACCESS METHOD FOR THE INQUIRY/RESPONSE CHANNEL, J.P.S. Bhullar and J. B. Anderson (Department of Electrical Engineering and Communications Research Laboratory, McMaster University, Hamilton, Ontario, Canada, L8S 4L7). In an inquiry/response channel, very many users send infrequent, short messages, and addressing information makes up a major component of the information flow. We propose a scheme to accomplish optimal multiple access for such a channel when the channel is a digital loop, and give a simple implementation. More than one user may successfully signal at once; without this feature, a system suffers either a loss of efficiency or a throughput delay. In the scheme's operation, each user in turn views a single circulating information word as it passes through him. If he wishes to signal a (short) message, he multiplies the word by one of a collection of words identified only with him. For information flowing the opposite way, the user checks to see if one of his words is a factor of the circulating word. For 1-bit messages, the prime numbers provide a good source of codes. By the prime number theorem, W words can be signalled up to N at a time

in a N log W bit word. Conversely, it can be shown that the message, addressing, and routing information requirement tends to about this many bits for small N. However, simpler schemes than our multiplication method exist which do this well for small N, so we stress instead the case for large N. Results are presented of searches for large-N codes, not necessarily based on primes, which almost always have an unambiguous decoding. Applications to polling and to the distribution of short messages are discussed. Encoder and decoder hardware requires a multiplier, divider, and a table look-up only.

SESSION F4

NEW BOUNDS TO R(D) OF A SOURCE WHOSE OUTPUT IS THE SUM OF TWO INDEPENDENT RANDOM ENTITIES, Dimitris Anastassiou and David J. Sakrison (Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, CA 94720). In an effort to apply results of information theory to the efficient storage or transmission of images, it has been found that the intensity random field of a typical image could be represented as the sum of two independent random fields of simpler description; it has also been found that, under some conditions, the frequency-weighted squared-error criterion is a satisfactory model for the fidelity evaluation of the images. Thus, a situation of practical interest is when the source output is the sum of two independent random entities, whose corresponding rate-distortion functions with respect to the (probably frequency-weighted) squared-error criterion, can be found or approximated. The values of these functions have been used to provide new bounds to the rate-distortion function of the additive source with respect to the same criterion. These bounds have been derived by use of either special encoding strategies or conditional rate-distortion theory; some of them are shown to be tighter than the known bounds for certain regions of distortion, and have a desirable asymptotic behavior for small distortion levels.

RATE-DISTORTION FUNCTIONS FOR CONTINUOUS ALPHABET MEMORYLESS SOURCES, Stephen L. Fix and David L. Neuhoff (University of Michigan, Ann Arbor, MI 48109). The analytical evaluation of the rate-distortion function for memoryless sources with continuous alphabets has been until recently limited to Gaussian sources and the squared error distortion criterion and cases where the Shannon lower bound is tight. Recently Tan and Yao have calculated the rate-distortion function for a Gaussian source with a magnitude-error distortion criterion and have shown that their procedure can be applied to similar sources. This talk deals with various analytical techniques available in finding rate-distortion functions. Specifically the squared-error distortion criterion will be discussed.

It has been thought that Blahut's algorithm would be an efficient way to compute R(D) for continuous alphabet sources. However, as will be discussed, for most sources and the squared-error distortion criterion the solution will be discrete and therefore the algorithm will need modification to work efficiently. A modified algorithm is suggested and examples discussed.

ON THE ASYMPTOTIC BEHAVIOUR OF THE RATE-DISTORTION FUNCTION, Gérard Cohen (ENST, 46 rue Barrault, 75013 Paris, France) and Claudia Lidia Simionescu (Department of Mathematics, University of Brasov, 2200 Brasov-R.S. Romania). The paper deals with a result on the asymptotic behaviour of the rate-distortion function for a quite general class of source distributions and difference distortion measures.

SESSION F4

AN APPROXIMATION TO RATE-DISTORTION, Gérard Cohen (Ecole Nationale Superieure des PTT, 46 Rue Barrault, 75013 Paris, France) and Claudia Lidia Simionescu (Department of Mathematics, University of Brasov, 2200 Brasov-R.S. Romania). The paper deals with a new approach to the problem of computing the rate-distortion function associated with a given source and a given fidelity criterion. For any $\delta>0$, we construct a minimal set of conditional probabilities $A_Q = \{Q_i\}$.

Minimization over A_Q gives an approximation to the rate-distortion function. A convergent algorithm for getting minimal sets A_Q is shown and implementation for the results has been done.

The method admits generalizations to continuous sources and more general distortion measures.

BOUNDS ON THE PERFORMANCE OF LINEAR SOURCE CODING, T. C. Ancheta (IBM Research, Yorktown Heights, NY). It is shown that when a binary stationary source with entropy H is transformed linearly into the compressed sequence utilizing $R_c \leq H$ output digits per source letter, the average Hamming distortion in reconstructing the source cannot be less than $(1-R_c)h^{-1}((H-R_c)/(1-R_c))$ where h is the binary entropy function. This bound lies above the rate-distortion function of the binary memoryless source (except at the points $R_c = \{H,0\}$ where they meet). For the binary symmetric source, the bound is readily shown to be tight.

We consider a universal generalization of linear source coding in which the source sequences are partitioned into M subsets and a linear mapping is associated with each subset. A source sequence is encoded by transmitting a $\log_2 M$-bit prefix to indicate the subset to which the sequence belongs followed by the compressed digits resulting from the mapping corresponding to such a subset. It is shown that the performance of this source coding scheme in compressing a stationary source is bounded aymptotically by the same function as above when the number of partitions is finite.

JOINT SOURCE-CHANNEL TIME-INVARIANT TRELLIS ENCODING, James G. Dunham and Robert M. Gray (Department of Electrical Engineering, Stanford University, Stanford, CA 94305). One of the fundamental problems of information theory is the joint source-channel coding problem, that is, the optimal transmission of a source over a noisy channel to a user with respect to a fidelity criterion. We show how the theory of sliding-block codes can be used to obtain a theory of joint source-channel time-invariant trellis encoding for discrete ergodic sources, discrete memoryless channels and bounded single-letter fidelity criteria. A block trellis encoder joint source-channel coding theorem is stated which shows that the Optimal

Performance Theoretically Attainable (OPTA) for these codes is given by D(C), the distortion-rate function D(·) evaluated at the channel capacity C. This generalizes the theory of source time-invariant trellis encoding from discrete memoryless noiseless channels to discrete memoryless noisy channels. Also, this leads to a new method of channel coding using a per symbol error criterion and these new channel codes are compared with time-invariant convolutional block channel codes.

TREE SEQUENTIAL ENCODING FOR SOURCES WITH MEMORY UNDER A FIDELITY CRITERION, Harry H. Tan (Department of Electrical Engineering and Computer Science, Princeton University, Princeton, NJ 08540). A sequential algorithm for encoding the outputs of discrete time stationary block ergodic sources under a single letter fidelity criterion using tree codes is analyzed. It is shown that for encoding these source-user pairs, there are tree codes that satisfy the rate-distortion bound with the algorithm used as the encoder. Moreover the average number of computations per encoded tree branch required by the algorithm is finite. Bounds on computational requirements are also developed.

CAUSAL RATE-DISTORTION THEORY FOR REAL-TIME ESTIMATION, Julian L. Center, Jr. (The Analytic Sciences Corporation, Six Jacob Way, Reading, MA 01867). Real-time estimation requires a causal estimation scheme; that is, future measurements cannot be used to construct current estimates. This paper extends rate-distortion theory to include causality constraints. A formulation of the rate-distortion problem with causality constraints is presented, and the character of the solution to this problem is discussed. For the linear-system, quadratic distortion, Gaussian-noise case, a technique yielding approximate solutions is presented. This technique leads directly to an easily-realizable system operating on the approximate rate-distortion boundary.

OPTIMIZING RECEIVERS FOR REMOTE PERCEPTION USING SENSORY NOISE REDUCTION TECHNIQUES, Charles Honorton (Department of Psychiatry, Maimonides Medical Center, Brooklyn, NY). We report a series of experiments demonstrating remote information transfer between a source (human sender + target message) and a sensorially-isolated human receiver in another room, 4.5 m distant. Receiver optimization techniques were used to enhance detection and retrieval of remote stimuli by attenuating patterned perceptual input and proprioception, considered as likely noise sources in remote perception. The generalizability of these findings is supported by independent replications in six different laboratories. A film demonstrating an experimental session will be shown preceding this talk.

AXIOMATIZATION AND REPRESENTATION OF QUALITATIVE INFORMATION, Zoltan Domotor (University of Pennsylvania). Given a Boolean field of measurable sets impressed with a binary comparative informativeness relation and a binary informational independence relation, necessary and sufficient conditions are provided for the existence of a Shannon information measure on the field such that it respects the independence and agrees with the comparative informativeness. The relevant axioms are tracked down by solving finite systems of bilinear inequalities in symmetric product spaces.

A representation using *non-Archimedean* information measures is presented, based on the compactness property of finite substructures.

The notion of information-preserving map is introduced for discussing some of the basic constructions, such as quotients, products, and filtered limits of information spaces.

Arguments are displayed in favor of qualitative information theory and some directions are suggested for future research.

INFORMATION THEORY AND ORGANIZATION THEORY, R. F. Drenick, (Polytechnic Institute of NY, 333 Jay Street, Brooklyn, NY 11201). The paper reports on a recent mathematical approach to organization theory which relies heavily on concepts drawn from information theory. These concepts were found to be useful in the modelling of the individual organization member, as well as in the characterization of the functioning of the organization as a whole. The model of the member in particular is based on Shannon's original definition of a noiseless channel. Extensions however were indicated in several directions, partly in order to accommodate observational evidence, and partly in order to define operations in the presence of "overload" which is not usually considered in information theory. It is shown that the desire to avoid overload in fact accounts for many of the most conspicuous features of organizations in practice. Concepts drawn from information theory can be used to describe their purpose and their performance.

SESSION F5

RECENT WORK ON ALGORITHMIC INFORMATION THEORY, Gregory J. Chaitin (IBM Research, Yorktown Heights, NY 10598). Algorithmic information theory is an attempt to apply information-theoretic and probabilistic ideas to recursive function theory. Typical concerns in this approach are, for example, the number of bits of information required to specify an algorithm, or the probability that a program whose bits are chosen by coin flipping produces a given output. During the past few years the definitions of algorithmic information theory have been reformulated. We review some of the recent work in this area.

COMPLEXITY BASED INDUCTION SYSTEMS: COMPARISONS AND CONVERGENCE THEOREMS, R. J. Solomonoff (Rockford Research, Inc., 140 1/2 Mt. Auburn Street, Cambridge, MA 02138). In 1964 we proposed as an explication of a priori probability, the probability measure induced on output strings by a universal Turing machine with unidirectional output tape and a randomly coded unidirectional input tape.

Levin has shown that if $\tilde{P}'_M(x)$ is the unnormalized form of this measure, and $P(x)$ is any computable probability measure on strings, x, then

$$\tilde{P}'_M(x) \geq C\, P(x)$$

C being a constant independent of x. We show that this result for P'_M, the normalized form of this measure, is directly derivable from Willis' probability measures on non-universal machines.

If the conditional probabilities of P'_M are used to approximate those of P, then the expected value of the total squared error in these conditional probabilities is bounded by $-1/2 \ln C$. In this error criterion, and when used as the basis of a universal gambling scheme, P'_M is superior to Cover's measure, b*.

When $H^* = -\log_2 P'_M$ is used to define the entropy of finite sequences, the equation

$$H^*(x,y) = H^*(x) + H^*_x(y)$$

holds exactly, in contrast to Chaitin's entropy definition, which has a non-vanishing error term in this equation.

SESSION G1

SYNDROME DECODING OF BINARY RATE-k/n CONVOLUTIONAL CODES, J. P. M. Schalkwijk, A. J. Vinck and K. A. Post (Eindhoven University of Technology, Eindhoven, The Netherlands). This paper concerns a state space approach to syndrome decoding of binary rate-k/n convolutional codes. State space symmetries of a certain class of codes can be exploited to obtain an exponential reduction of the hardware of a maximum likelihood decoder. Simulation results indicate that the state space symmetries can also be used to advantage in sequential decoding. In Fano and stack decoding, one obtains major savings in the number of computations and in the required storage. Aside from these practical advantages, it is felt that the state space formalism developed in this paper has some intrinsic value of its own.

CONCATENATED CODES FOR IMPROVED PERFORMANCE WITH APPLICATIONS TO THE RAYLEIGH FADING CHANNEL, John F. Pieper (Stein Associates, Waltham, MA), John G. Proakis (Northeastern University, Boston, MA), Roger R. Reed (Stein Associates, Waltham, MA), and Jack K. Wolf (University of Massachusetts, Amherst, MA). The concatenation of two block codes is an effective procedure for generating a block code of very large size. A specific form of concatenation and the associated decoding algorithm are briefly discussed. Performance results for a Rayleigh fading channel are presented which show that a savings of several dB is obtainable in this manner.

CONTINUED FRACTIONS AND BERLEKAMP'S ALGORITHM, L. R. Welch and R. A. Scholtz (Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90007). The problem of determining the error-locator polynomial and error-evaluator polynomial in decoding BCH codes is equivalent to determining a ratio of polynomials from a given number of coefficients in its formal power series expansion. Viewing polynomials as a ring of "integers" within a field of formal power series, the decoding problem can be translated into one of finding rational approximations to field elements.

Continued fractions have been used for many years to generate good rational approximations to real numbers. In this paper the theory of continued fraction approximations is developed abstractly for ring-field combinations on which a non-archimedean norm is defined and for which the "integer-part" mapping from field elements to ring elements satisfies two simple postulates. Basic theorems concerning the optimality of continued fraction approximations generated in this setting are given and the connection with the BCH decoding problem is established.

SESSION G1

A DECODING PROCEDURE FOR MULTIPLE-ERROR-CORRECTING CYCLIC AN CODES, Wen-Yung Yeh (Department of Electrical Engineering, Northeastern University, Boston, MA 02115). A decoding procedure for multiple-error-correcting arithmetic cyclic codes is proposed. For cyclic AN codes $C_r(n,B,t)$ with information rate $\log_r B/n$ less than $1/t$, the error trapping technique will correct all correctable error patterns. If the information rate is larger than $1/t$ but less than $2/t$, a modified error-trapping decoding scheme has been developed. Fortunately it turns out that all useful multiple-error-correcting cyclic AN codes with block length n less than 58, listed in the appendix B of a previous thesis by Larson, can be decoded effectively by this algorithm.

ARITHMETIC CODES IN RESIDUE NUMBER SYSTEMS, Ferruccio Barsi and Piero Maestrini (Istituto di Elaborazione dell'Informazione del CNR, Pisa, Italy). Residue Number Systems (RNS) provide a means to construct non-binary, multiple error correcting arithmetic codes. In this paper two different codes, namely systematic codes in RNS and AN codes in RNS are considered and their properties are discussed in detail. A lower bound to redundancy allowing t-correction in a class of codes including those under consideration is reported and it is shown that systematic codes and AN codes in RNS reach this bound. In both codes, error correction is achieved by finding appropriate solutions to a key congruence. Two different decoding procedures are presented. The first one, although working with any t-correcting code, is computationally inefficient for high values of t. The second procedure, based upon the euclidean algorithm for integers, is very efficient but requires an amount of redundancy slightly above the lower bound. The error correcting procedure based upon the euclidean algorithm was previously known for AN codes in RNS. In the case of systematic codes in RNS, this procedure is faster than the best previously known procedure, requiring multiple iterations of an algorithm based on convergence of continued fractions.

A NOTE ON THE IDENTIFICATION OF TWO-DIMENSIONAL ARMA MODELS, H. W. Penm and M. Kanefsky (University of Pittsburgh). Recent work has focused on modeling and enhancement of two-dimensional monochromatic images. A major complication in extending one-dimensional techniques is the fact that finite order factorization of the spectrum $\phi_o(Z_1,Z_2)$ is usually not possible. Woods has made progress in the factorization problem but truncation methods are necessary for finite recursive filtering techniques. Strintzis assumes that the power spectrum itself can be modeled by finite factors or

$$\phi_o(Z_1, Z_2) = \left(\sum_{i=0}^{M_1}\sum_{j=0}^{M_2} A_{ij}\, Z_1^i\, Z_2^j\right)^{-1} \left(\sum_{i=0}^{N_1}\sum_{i=0}^{N_2} f_{ij}\, Z_1^i\, Z_2^j\right),$$

where $N_1 \geq M_1$, $N_2 \geq M_2$.

This two-dimensional $(N_1, N_2; M_1, M_2)$ ARMA model results in optimal finite recursive filters which require the identification of the A_{ij} or AR coefficients.

A procedure for identifying this two-dimensional ARMA model (i.e., both the order and the AR coefficients) from the measured spatial correlation matrix is presented. This procedure is a modification of the one-dimensional technique of Graupe, Krause and Moore, and is easily extended to N-dimensional data.

While the ARMA model does not permit simulation of data, effects of measurement inaccuracies can still be analyzed. This can be accomplished by simulating measured spatial correlation function directly. Using such procedures, the effects of measurement error on an example were investigated and no unexpected instability problems were evident.

THE RECURSIVE IDENTIFICATION OF STOCHASTIC SYSTEMS USING AN AUTOMATON WITH SLOWLY GROWING MEMORY, B. D. Kurtz (Howden Applied Research, Toronto, Canada) and P. E. Caines (Division of Applied Sciences, Harvard University, Cambridge, MA). Consider a discrete time system or process with i.i.d. outputs, $X_1, X_2, \ldots$ The distribution of X_i's is characterized by some unknown parameter, θ, which is an element of Z^+, the positive integers. We describe an algorithm (automaton) which possesses a slowly growing memory and which generates estimates converging to θ w.p.1.

The following three specific applications are discussed:

1) Identification of the most probable value of a single sample from a set of discrete i.i.d. random variables.

2) Identification of the parameter of an observed sequence of i.i.d. Bernoulli random variables.

3) Identification of the parameters of a moving average process. (Notice that the observations of the system output in this case are not i.i.d.).

An estimate of the rate of convergence of the algorithm exceeds that of other proposed recursive algorithms for parameter identification. Furthermore, the amount of memory required is shown to grow as log log i, a rate significantly slower than that required by other algorithms. It is also shown that for a very large class of systems, convergence of parameter estimates is impossible without access to an unbounded memory.

Simulation experiments are described.

MLE OF DENSITY FORMS AND IDENTITY OF NORMAL MIXTURES, D. S. Arantes (Universidade Estadual de Campinas, Caixa Postal 1170, Campinas - S.P. - Brazil). Given a set of i.i.d. univariate random variables $x_1, x_2, x_3, \ldots, x_n$ distributed according to f(x), the problem is to estimate the density f(x) or a set of parameters θ_o characterizing f(x). Strong convergence in information (divergence) is established for maximum likelihood estimates of densities under hypotheses similar to, but distinct from, those of Wald. The results are then applied to the problem of estimating the parameters characterizing a family of countable mixtures of normal densities.

CONSISTENT ESTIMATION OF THE ORDER OF AUTOREGRESSIVE SYSTEMS, William G. Hwang (School of Electrical Engineering, Cornell University). Hitherto suggested estimators of system order have not been consistent. We propose a statistically consistent, but computationally demanding, estimator of the order of an autoregressive process. We also present results on the efficiency of our method and the best possible efficiency of any estimator of system order.

RESOLUTION ENHANCEMENT IN THE AUTOREGRESSIVE SPECTRAL ESTIMATION OF NOISY SIGNALS, M. Kaveh (Department of Electrical Engineering, University of Minnesota, Minneapolis, MN 55455). The autoregressive (AR) method is becoming a preferred mode of spectral estimation. A main advantage of this technique is that it is data-adaptive and free from the effects of window functions associated with the traditional Blackman and Tukey type methods. A main weakness of the AR spectral estimation method is the strong dependence of its resolution on the signal-to-noise ratio of the data.

In this paper a modified autoregressive model based on the autocorrelation samples for lags beyond the inverse noise bandwidth is followed by an all-zero model estimated from the autocorrelation samples in the inverse of the noise bandwidth. For small signal-to-noise ratios this model requires many fewer parameters than the pure autoregressive one in estimating the power spectrum of signals with very high resolution. Some examples are shown that compare this method with the pure AR and the optimum autoregressive-moving average spectral estimators.

SESSION G3

PSEUDO-RANDOM NUMBERS, H. G. Niederreiter (Mathematics Department, University of Illinois, Urbana, IL 61801). Pseudo-random numbers are used for the deterministic simulation of random variables and random processes. The most commonly employed pseudo-random numbers for simulating the uniform distribution are those generated by the linear congruential method. The main part of the talk will be devoted to recent progress concerning the statistical properties of such pseudo-random numbers.

The performance of linear congruential pseudo-random numbers under a variety of statistical tests has been investigated. An important requirement for the applicability of pseudo-random numbers in Monte Carlo calculations is statistical independence of successive terms. Effective results on the deviation of linear congruential pseudo-random numbers from an ideal independence behavior are now available. The results lead to criteria for the optimum choice of parameters in the linear congruential method in order to produce a desired independence property. Information about pseudo-random numbers generated by feedback shift registers will also be given.

SOME ASPECTS OF CONVEXITY THAT IMPACT INFORMATION THEORY, Hans Witsenhausen (Bell Telephone Laboratories, Murray Hill, NJ 07974). From its very beginning, information theory was pervaded by convexity arguments. Much of the necessary background was developed on an ad hoc basis without reference to the knowledge existing in the mathematical study of convex sets and functions. Yet, explicit use of this knowledge is helpful in several areas of information theory, as will be shown by examples.

RECENT RESULTS IN ERGODIC THEORY, Paul C. Shields, (Mathematics Department, University of Toledo, Toledo, OH 43606). Several equivalent ways of deciding whether or not a stochastic process is a sliding-block coding of an i.i.d. process will be described. In the course of these descriptions such concepts as the $\bar{d}$-distance between processes, Rohlin stacks, and the Ornstein isomorphism theorem will be discussed.

MARKOV RANDOM FIELDS, Frank L. Spitzer (Mathematics Department, Cornell University, Ithaca, NY 14853). Markov random fields are a special class of binary stationary stochastic processes whose parameters take values in Z_d, the d-dimensional lattice points. Discussion will focus on four equivalent ways of characterizing a Markov random field, namely:

1) in terms of conditional probabilities depending only on nearest neighbor

2) as a Gibbs state with nearest neighbor potential

3) as a time reversible equilibrium state of certain time evolutions with nearest neighbor interaction

4) as a state with lowest free energy or highest entropy, among all stationary states with the same given specific energy.

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX

AUTHOR INDEX